The Dissertation Committee for Yeonjin Ko Certifies that this is the approved version of the following dissertation:

Unusual Carbohydrate Biosynthesis: Mechanistic Studies of DesII and the Biosynthesis of Formycin A

Committee:

Hung-wen Liu, Supervisor

Eric V. Anslyn

Adrian T. Keatinge-Clay

Christian P. Whitman

Walter L. Fast

Unusual Carbohydrate Biosynthesis: Mechanistic Studies of DesII and the Biosynthesis of Formycin A

by

Yeonjin Ko, B.S.CH.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2017

Acknowledgements

Joining the Liu Group and studying enzyme mechanisms have been extremely rewarding experiences in my life. It would have not been even possible to complete such a long and challenging journey without the many people around me.

First of all, I would like to thank my research advisor, Professor Hung-wen Liu. He always supports me and gives great advice not only on the scientific logics and experiments but also on the scholarly attitude. I am honored to be one of his students, and I would like to follow his example of enthusiasm for science. I also want to thank my dissertation committee members, Drs. Eric Anslyn, Adrian Keatinge-Clay, Christian Whitman, and Walter Fast, for their guidance.

I feel grateful to every member in the Liu Group. Especially I would like to extend my appreciation to Dr. Mark Ruszczycky for helping me to design and conduct experiments and to analyze data. He also helped me prepare the dissertation. I thank Dr. Yasushi Ogasawara for his excellent input into genome analysis and identification of gene cluster. Dr. Sei-Hyun Choi, Geng-Min Lin and Shao-An Wang are experts in synthesis, and they made great contributions to the projects. I am also grateful to Dr. Yung-Nan Liu who is the best resource in molecular biology. I have really enjoyed being a member of the Liu Group, and I am thankful that I have the chance to work with such a wonderful group of people.

I would like to give special thanks to my better half, Byung Joon, for his love and support along the way, and to my two-year-old boy, Sean, who is a true blessing in my life. I greatly appreciate the strong support and prayer by my parents on the other side of the world. Finally, I thank God and believe in His perfect plan for me.

Unusual Carbohydrate Biosynthesis: Mechanistic Studies of DesII and the Biosynthesis of Formycin A

Yeonjin Ko, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Hung-wen Liu

Carbohydrates are essential biomolecules in all living organisms. Besides serving as energy storage and structural building blocks in primary metabolism, carbohydrates represent the building blocks for numerous bioactive natural products. The presence of sugar moieties in secondary metabolites are important for the biological properties of these natural products. Thus, the study of biosynthetic pathways involving carbohydrate secondary metabolism can reveal some intriguing enzymatic transformations that lead to remarkable structural diversity. Understanding these pathways and the enzymes they contain can provide new insights for pathway engineering and the production of novel sugar structures.

The radical *S*-adenosyl-L-methionine (SAM) enzymes are distinguished by their unique chemistry that results in the reductive homolysis of SAM to generate a reactive 5′-deoxyadenosyl radical. Subsequent formation of a substrate radical intermediate by this radical initiator permits a diverse set of biotransformations. DesII belongs to the radical SAM enzyme superfamily, and is involved in the biosynthesis of TDP-desosamine in *Streptomyces venezuelae*. DesII catalyzes the deamination of its biosynthetic substrate (*i.e.*, TDP-4,6-dideoxy-3-keto-D-glucose), whereas it promotes an oxidative dehydrogenation reaction when the C4 amino group of the substrate is replaced by a hydroxyl group (*i.e.*,

TDP-D-quinovose). Control of the radical intermediate resulting in two distinct reaction outcomes has been a primary focus of this research. It has been proposed that the orbital geometry of the radical intermediates is an important factor in determining whether the enzyme functions as a lyase or a dehydrogenase. To investigate this hypothesis, several substrate analogs with altered stereochemistry of ring substituents were tested as potential substrates for DesII and the reaction products were characterized. While DesII deaminates the C4 axial amino substituent as it does with the C4 equatorial amino group, inversion of stereochemistry of a hydroxyl group at C4 allows dehydration to take place. These results support the working hypothesis that the stereochemical configuration of substrate radical in the active site plays an important role in controlling the partitioning into different reaction pathways.

Formycin A and coformycin are nucleoside antibiotics produced by *Nocardia interforma* and *Streptomyces kaniharaensis*. Their biosynthetic pathways are of particular interest because of their unusual structural features such as the pyrazolopyrimidine nucleobase with a *C*-glycosidic linkage in formycin A and the 1,3-diazepine ring in coformycin. Genomic analysis of producing strain suggested that the pyrimidine ring of formycin A is formed in a pathway analogous to that for purines, which led to the identification of a potential biosynthetic gene cluster and pathway for formycin A. Conversion of the putative intermediate, carboxyaminopyrazole ribonucleotide, in this pathway to the final product was demonstrated to be catalyzed by enzymes encoded in the *for* cluster as proposed. It was also shown that one gene adjacent to the formycin A gene cluster encodes a reductase that catalyzes the last step in the biosynthesis of coformycin. This study aims to elucidate the biosynthetic pathways of formycin A and coformycin with an emphasis on the formation of pyrazolopyrimidine moiety and *C*-glycosidic bond.

Table of Contents

List of Tables

List of Figures

Chapter 1: Background and Significance

1.1 Radical SAM Enzymes in the Biosynthesis of Carbohydrates

1.1.1 Radical SAM Enzymes

Radical *S*-adenosyl-L-methionine (SAM) enzymes represent one of the most rapidly growing enzyme superfamilies found across all kingdom.[1-3] Radical SAM enzymes utilize a [4Fe–4S] cluster to initiate a reductive homolytic cleavage of SAM (**1**) that generates a 5′-deoxyadenosyl radical (**2**, 5′-dAdo•) and methionine (**3**) (Figure 1-1). The 5′-deoxyadenosyl radical (**2**) subsequently abstracts a hydrogen from the substrate to produce 5′-deoxyadenosine (**4**, 5′-dAdo) and a substrate radical (denoted R• in Figure 1-1) that can undergo a variety of radical-mediated transformations. As the hydrogen atom can be abstracted from an unactivated C–H bond, these enzymes are capable of catalyzing a diverse set of reactions that would be challenging to accomplish via Lewis acid/base chemistry alone. Radical SAM enzymes typically contain three conserved cysteine residues coordinated to the [4Fe–4S] cluster, usually in a characteristic CX_3CX_2C motif, although the number of residues between cysteines can be varied. The reactions catalyzed by radical SAM enzymes include but are not limited to methylation, oxidation, sulfur insertion, and complex rearrangements. This class of enzymes have been extensively reviewed elsewhere.[3]

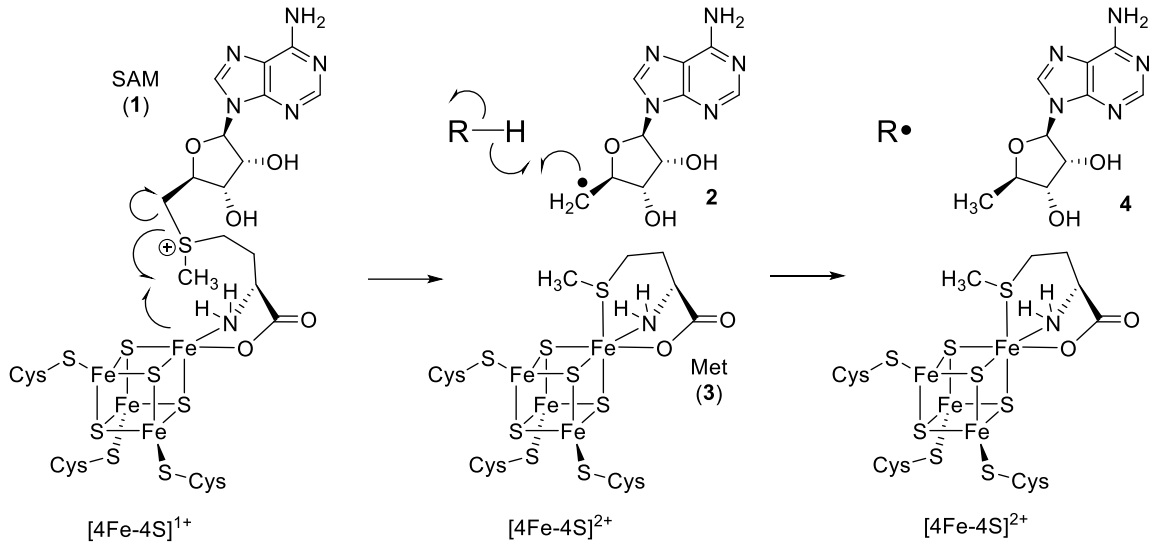


Figure 1-1. Reductive homolysis of SAM (**1**) during the catalytic cycle of radical SAM enzymes.

1.1.2 Radical-Mediated Reactions in the Biosynthesis of Carbohydrates

Carbohydrates play a vital role in all biological systems, with prominent roles in energy storage and the structure of nucleic acids. In addition to primary metabolism, sugar biosynthesis is also an essential part of secondary metabolism, because the biological activities of many antibacterial and antiproliferative natural products depend on the presence of carbohydrate appendages.[4] Recent studies have shown that radical SAM chemistry is often employed to achieve complex transformations during the biosynthesis of unusual carbohydrates in nature.[5] The utilization of radical SAM enzymes in the biosynthesis of glycosylated natural product contributes to the structural diversity of the sugar products resulting from radical-mediated reactions. Mechanistic study of these enzymes can therefore help to unravel the principles by which reactive radical intermediates are generated and controlled in order to effect catalysis. Moreover,

understanding the biosynthetic pathways of unusual carbohydrates would offer the opportunity to exploit their biosynthetic machinery to generate novel sugar structures with enhanced biological activities.

BtrN is a well characterized radical SAM enzyme which is involved in the biosynthetic pathway of butirosin (**10**). Butirosin is an aminoglycoside antibiotic produced by *Bacillus circulans*. It contains a 2-deoxystreptamine (**9**, DOS) aglycone core structure that is also found in kanamycin, tobramycin, and gentamicin.[6] The discovery of the biosynthetic gene cluster of butirosin had allowed the elucidation of DOS biosynthesis, and the results showed that the pathway involves a 2-deoxy-*scyllo*-inosose (**6**, DOI) synthase BtrC, an aminotransferase BtrS, and a dehydrogenase BtrN (Figure 1-2).[7,8] The radical SAM enzyme BtrN was shown to be responsible for the production of 3-amino-2,3-dideoxy-*scyllo*-inosose (**8**, amino-DOI) from 2-deoxy-*scyllo*-inosamine (**7**, DOIA).[9] The dehydrogenation reaction catalyzed by BtrN is distinguished from the nicotinamide-dependent dehydrogenation reaction commonly found in the biosynthesis of other DOS-containing aminoglycosides, as catalysis by BtrN involves radical intermediates.

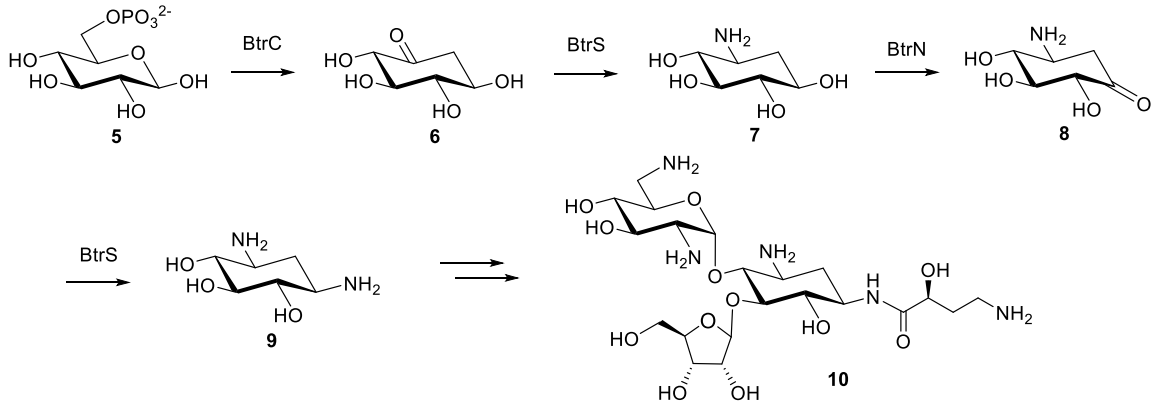


Figure 1-2. Biosynthetic pathway of butirosin (**10**).

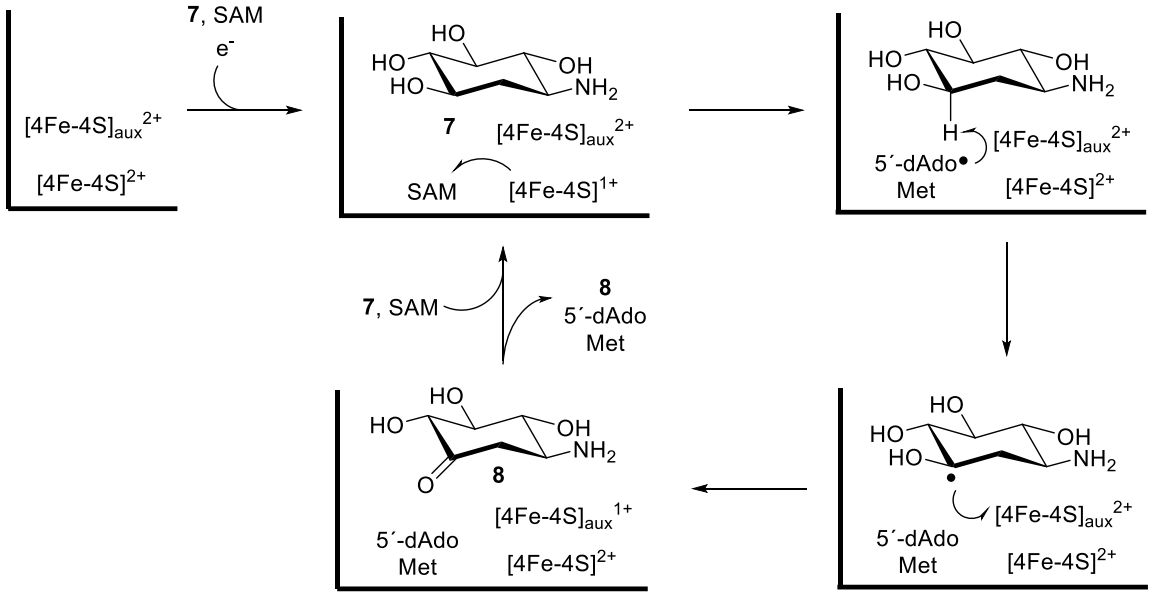


Figure 1-3. Proposed mechanism of BtrN-catalyzed dehydrogenation reaction of **7**.

The reaction mechanism of BtrN has been studied *in vitro* by biochemical characterization and electron paramagnetic resonance (EPR) spectroscopy (Figure 1-3).[10] The reaction is initiated by electron transfer from the reduced $[4Fe–4S]^{1+}$ cluster to SAM upon binding of SAM and the substrate DOIA (**7**) in the enzyme active site. Homolysis of the C–S bond of SAM generates 5′-deoxyadenosyl radical (**2**) that abstracts a hydrogen at C3 of **7**. Following the formation of the substrate radical, the dehydrogenation proceeds via one electron transfer back to the oxidized $[4Fe–4S]^{2+}$ cluster to convert the hydroxyl group at C3 to the corresponding keto group. Iron/sulfur titration, mutation, Mössbauer spectroscopy studies[11] along with X-ray crystallography[12] have revealed the presence of an additional auxiliary [4Fe–4S] cluster in the BtrN active site. Direct electrochemical measurement of redox potentials for the [4Fe–4S] clusters have also shown that the auxiliary cluster exhibits an unusually low potential.[13] While the role of the second cluster

is uncertain, it may participate in catalysis as an electron acceptor during oxidation of the substrate radical (see Figure 1-3).

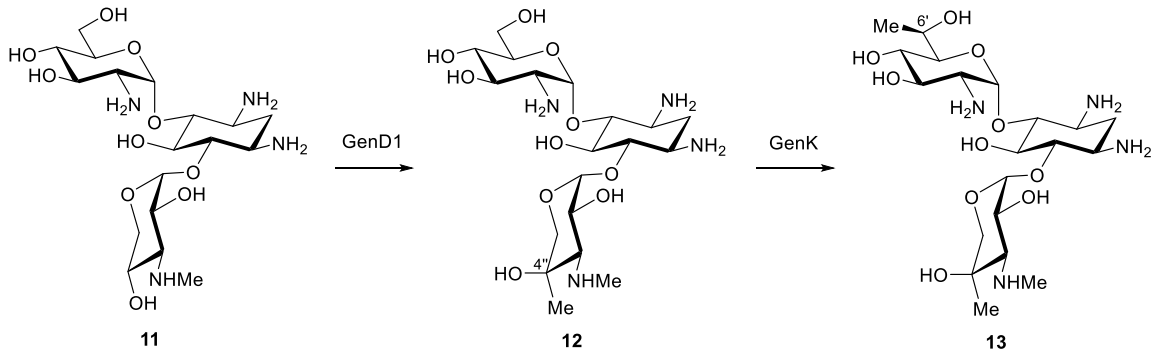


Figure 1-4. Methylation by GenD1 and GenK.

Some radical SAM enzymes require cobalamin as an additional cofactor.[14] GenD1 and GenK are cobalamin-dependent radical SAM enzymes involved in the biosynthesis of gentamicins in *Micromonospora echinospora*.[15] Gentamicins are clinically important aminoglycoside antibiotics that inhibit protein synthesis and are used in the treatment of Gram-negative bacterial infections.[6] During the biogenesis of gentamicins, GenD1[16] and GenK[17] catalyze the methylation of C-4″ of gentamicin A (**11**) and C-6′ of gentamicin X_2 (**12**), respectively (Figure 1-4). While GenD1 can be expressed in *Eschericha coli* as a soluble protein,[16] GenK was isolated as inclusion bodies and must be subjected to a series of denaturation and renaturation to make it soluble prior to reconstitution of its [4Fe–4S] cluster.[17] Both enzymes methylate unactivated sp^3 carbons in the presence of SAM, reductant, and cobalamin. NADPH together with methyl viologen acting as a redox

mediator was found to be a reliable source of reducing equivalents supporting the catalysis *in vitro*. The formation of G-418 (**13**) by GenK was accompanied by the production of a stoichiometric amount of 5′-dAdo (**4**) and *S*-adenosylhomocysteine (**14**, SAH), indicative of the consumption of two equivalents of SAM per turnover. After the initial hydrogen atom abstraction from C-6′ of gentamicin X_2 (**12**), a methyl radical can be transferred from methylcobalamin to the substrate radical to yield the product and cob(II)alamin (Figure 1-5, pathway A). Methylcobalamin may then be regenerated by reduction and methyl transfer from SAM. Alternatively, the formation of a substrate radical followed by deprotonation to afford a ketyl radical enables the nucleophilic methyl transfer reaction that generates cob(I)alamin (Figure 1-5, pathway B). The resulting alkoxy product radical should be reduced. The conversion of gentamicin A to gentamicin X_2 by GenD1 also generated 5′-dAdo, and the reaction mechanism of GenD1 is thought to be similar to that of GenK.

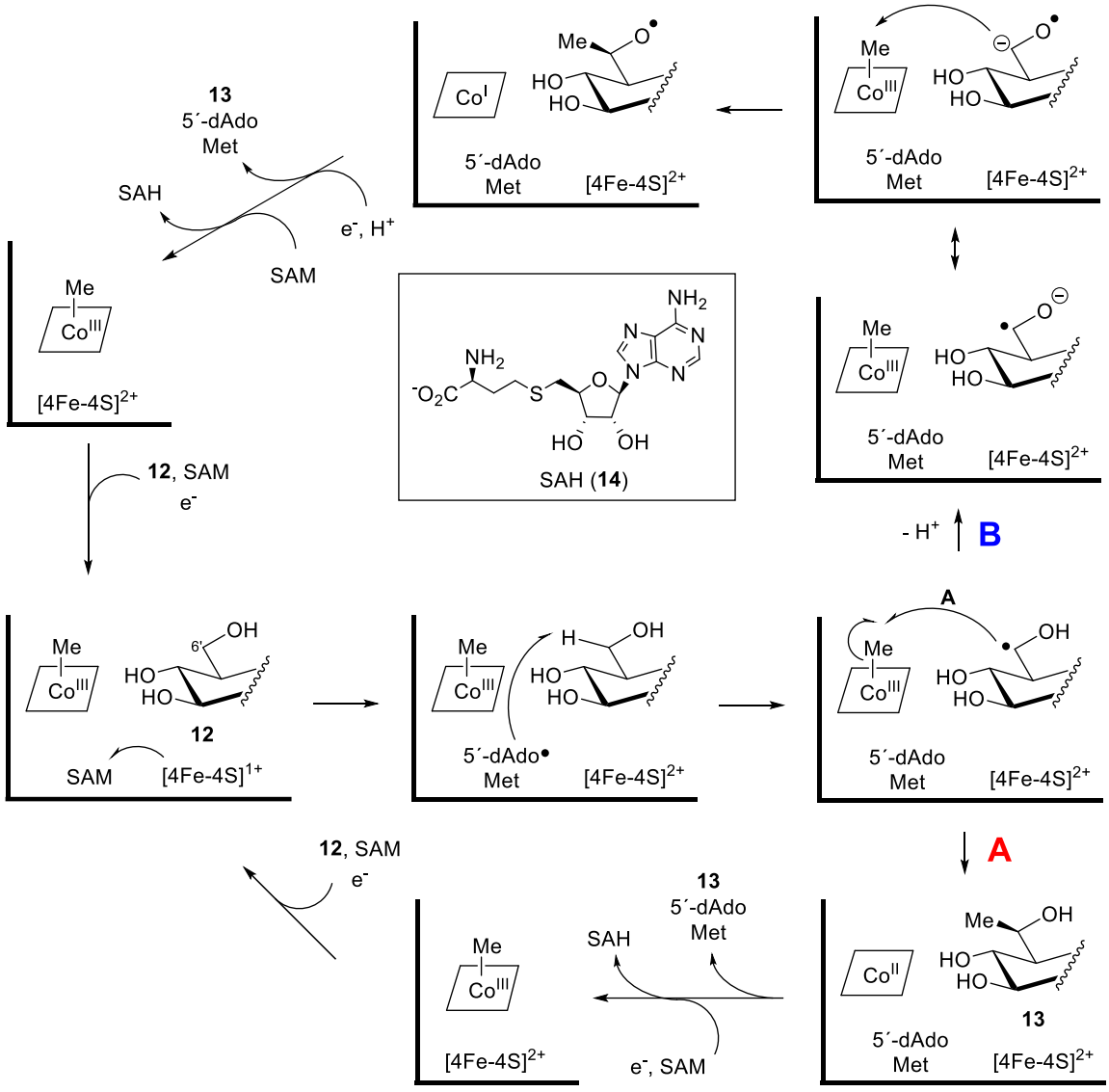


Figure 1-5. Proposed mechanism of methylation by GenK.

Similarly, a cobalamin-dependent radical SAM methyltransferase Fms7 was proposed to catalyze the *C*-methylation of fortimicin KL1 (**15**) during the biosynthesis of fortimicin A (**17**), an aminoglycoside antibiotic produced by *Micromonospora olivasterospora*.[18] MoeK5 was also implicated as a cobalamin-dependent radical SAM methyltransferase involved in the biosynthesis of moenomycin (**20**) in *Streptomyces ghanaensis*.[19] Deletion of the *moeK5* gene led to accumulation of the desmethylated derivative of moenomycin (**18**), implying that MoeK5 is required for the introduction of methyl group at C4 of the F ring.

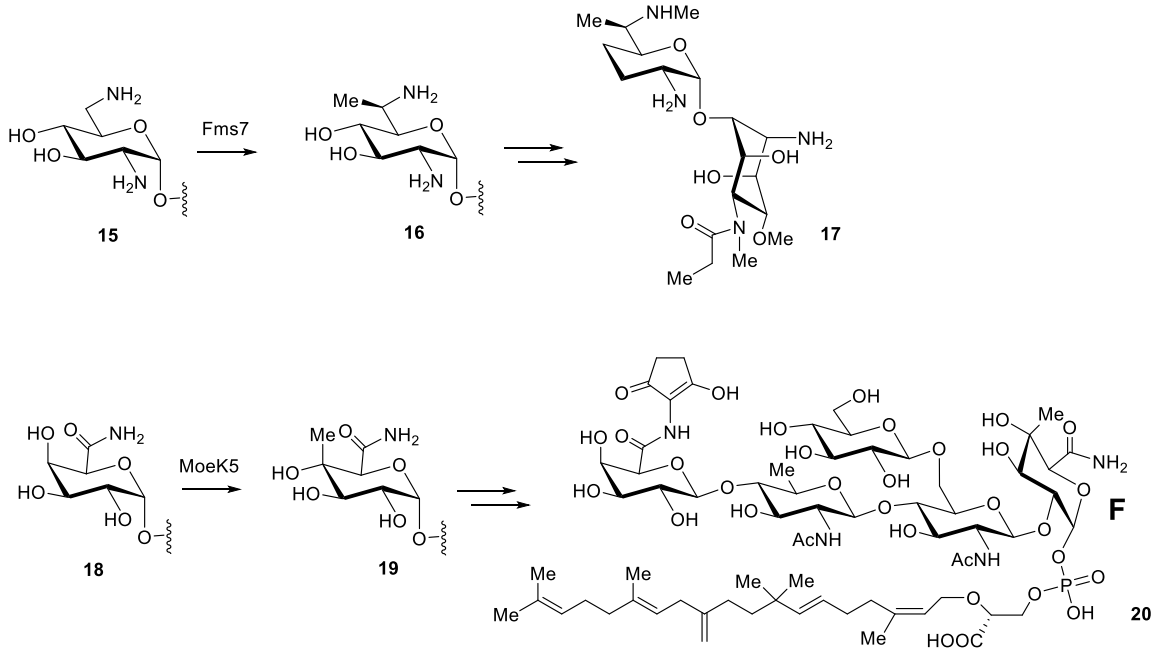


Figure 1-6. Fms7 and MoeK5 in the biosynthesis of fortimicin A (**17**) and moenomycin (**20**), respectively.

Oxetanocin A (**23**) produced by *Bacillus megaterium* is a nucleoside antibiotic with an unusual four-membered oxetane ring.[20] The biosynthetic gene cluster for oxetanocin is encoded on a plasmid and harbors the putative radical SAM enzyme OxsB.[21] OxsB catalyzes the conversion of deoxyadenosine nucleotide (**21**) to oxetane aldehyde (**22**) in the presence of SAM, cobalamin, and a reducing system.[22] The activity of OxsB also requires OxsA, which is a phosphatase encoded in the gene cluster (Figure 1-7). The detailed mechanism regarding the roles of cobalamin and OxsA during the ring contraction reaction is actively being investigated.

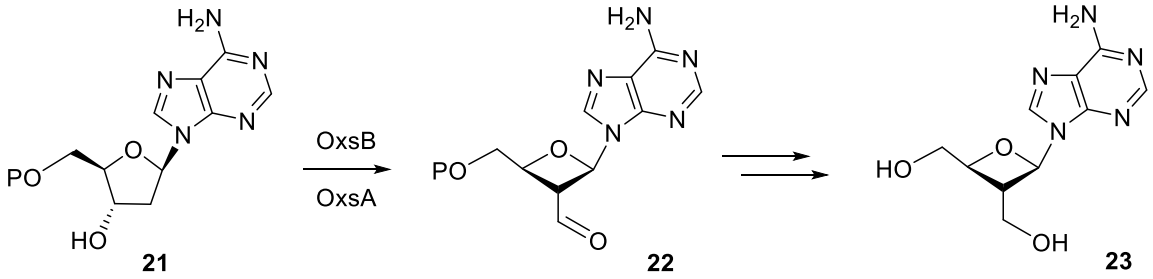


Figure 1-7. Ring contraction reaction catalyzed by OxsB during the biosynthesis of oxetanosine A (**23**).

The tunicamycins (**30**) are fatty acyl nucleoside antibiotics targeting bacterial cell wall biosynthesis and eukaryotic *N*-glycoprotein synthesis.[23] Their structures feature a unique 11-carbon core, tunicamine, and appended with fatty acyl chains.[24] Heterologous expression of the putative biosynthetic gene cluster from *Streptomyces chartreusis* identified the minimal gene set essential for the production of tunicamycin, and a likely pathway has been proposed based on bioinformatics analysis and gene disruption experiments (Figure 1-8).[25,26] The key finding in this pathway is the unusual *exo*-glycal intermediates.[27] TunA converts UDP-*N*-acetylglucosamine (**24**, UDP-GlcNAc) to UDP-6-deoxy-5,6-ene-GlcNAc (**25**), and then TunF epimerizes this compound to generate a galactose derivative (**26**, UDP-6-deoxy-5,6-ene-GalNAc). TunB is annotated as a radical SAM enzyme, and it is expected to be responsible for the coupling reaction between **26** and a uridyl moiety (**28**). The hydrogen abstraction from C-5′ of **28** would lead to radical addition to the exo-glycal intermediate (**26**) to form a stabilized radical intermediate. It was observed that the *exo*-glycal intermediate accumulated in the *tunB* deletion mutant, suggesting the essential role of TunB in the radical-mediated C–C bond formation. TunM is predicted to be a radical SAM methyltransferase, although its involvement has not been clearly demonstrated.

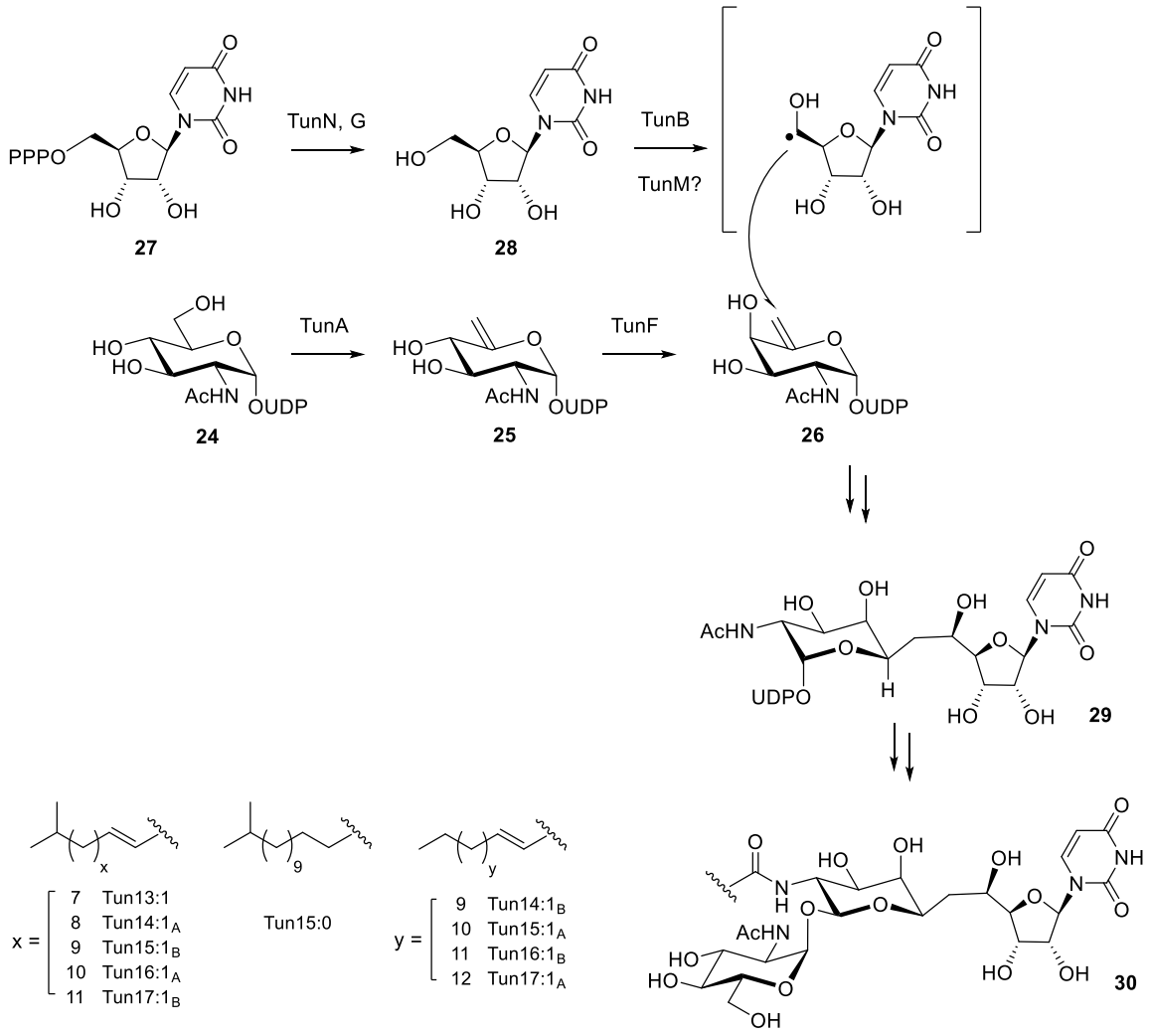


Figure 1-8. Biosynthesis of tunicamycins (**30**).

Blasticidin S (**34**)[28] and mildiomycin (**35**),[29] produced by *Streptomyces griseochromogenes* and *Streptoverticilium rimofaciens*, respectively, are representative peptidyl nucleosides that inhibit protein synthesis by binding to the large ribosomal subunit (Figure 1-9). Given that they share a common cytosyl pyranoside core structure, recent bioinformatics studies revealed that several gene products found in their biosynthetic gene clusters, which include putative radical SAM enzymes BlsE and MilG, are indeed homologous to each other.[30] The reconstituted BlsE was shown to catalyze the decarboxylation of cytosylglucuronic acid (**31**, CGA) to produce cytosylarabinopyranose (**33**) in the presence of dithionite or flavodoxin/flavodoxin reductase.[31] The disruption of *milG* resulted in the accumulation of 5-hydroxymethyl-CGA (**32**) implying that the formation of 2,3-dideoxy-2,3-ene-pyranose moiety is related to the function of MilG.[30] However, the mechanistic details of those reactions require further characterization.

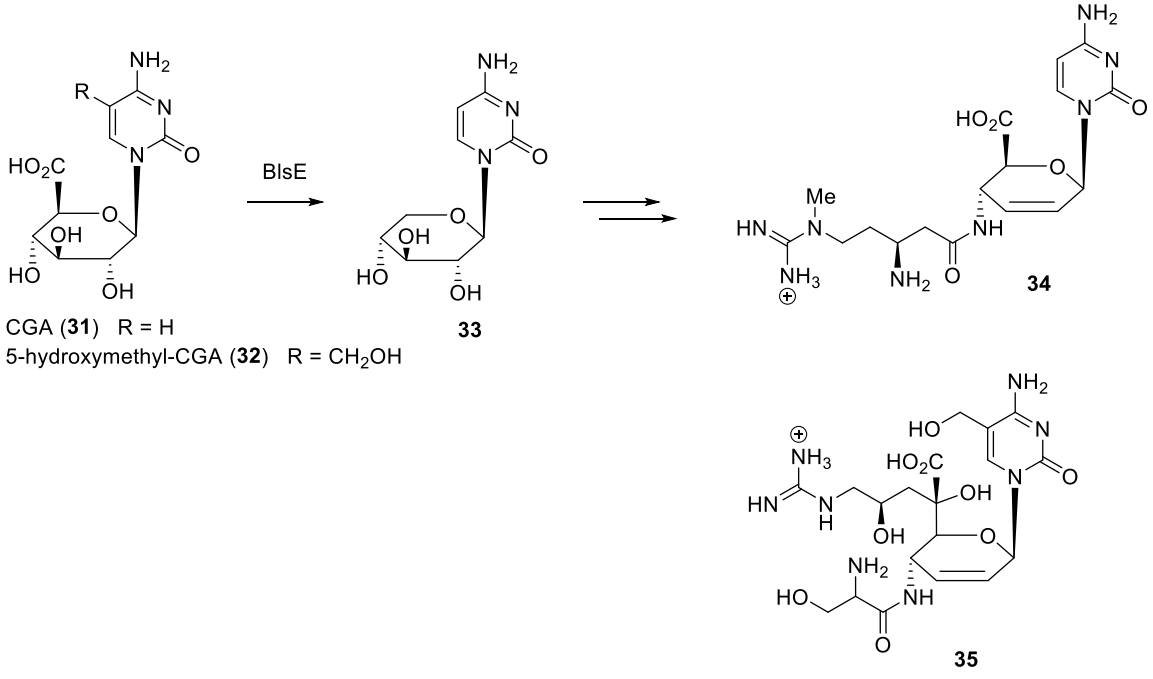


Figure 1-9. BlsE-catalyzed decarboxylation in blasticidin S (**34**) biosynthesis and the structure of mildiomycin (**35**).

Avilamycins (**36**) isolated from *Streptomyces viridochromogenes* are oligosaccharide antibiotics active against Gram-positive bacteria.[32] Structurally, avilamycins consist of a dichloroisoeverninic acid moiety with a heptasaccharide moiety (Figure 1-10). The biosynthetic gene cluster harboring 54 open reading frames was identified and contains a putative radical SAM enzyme *aviX12*.[33] The *aviX12* inactivated mutant produced a new derivative named gavibamycin N1 (**37**), a modified heptasaccharide with a glucose moiety in place of the mannose moiety.[34] Thus, AviX12 is proposed to catalyze the epimerization reaction during the biosynthesis of avilamycin, a reaction necessary to its biological activity. Currently, AviX12 has yet to be characterized *in vitro* so that further investigation will be necessary to establish AviX12 as a radical SAM epimerase.

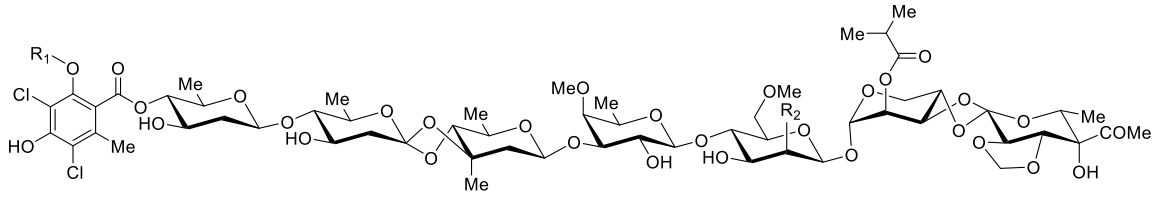


Figure 1-10. Structure of avilamycin A (**36**) and gavibamycin N1 (**37**).

1.1.3 DesII in the Biosynthesis of Desosamine

Desosamine (*i.e.* 3-(dimethylamino)-3,4,6-trideoxyhexose), is an essential structural component in many macrolide antibiotics such as erythromycin, methymycin, narbomycin, and pikromycin.[35] The binding of macrolides to the ribosomal subunit is mediated by several hydrogen-bonding interactions which involve the deoxysugar moiety. Identification of the biosynthetic gene cluster for desosamine in *Streptomyces venezuelae* revealed the presence of eight open reading frames in the cluster,[36] and one of the gene products, DesII, has been shown to be a radical SAM enzyme responsible for the C4 deoxygenation step (Figure 1-11).[37] Derived from TDP-D-glucose (**38**) by DesIV, conversion of TDP-4-keto-6-deoxy-D-glucose (**39**) to the corresponding C4 aminosugar (**40**) is catalyzed by the PLP-dependent aminotransferase DesI. Subsequent removal of the amino group at C4 is catalyzed by the radical SAM enzyme DesII.[38] Transamination and *N*,*N*-dimethylation steps are catalyzed by DesV and DesVI, respectively, to generate TDP-desosamine (**43**). The following glycosylation to modify the macrolactone ring is mediated by DesVII and DesVIII.[39]

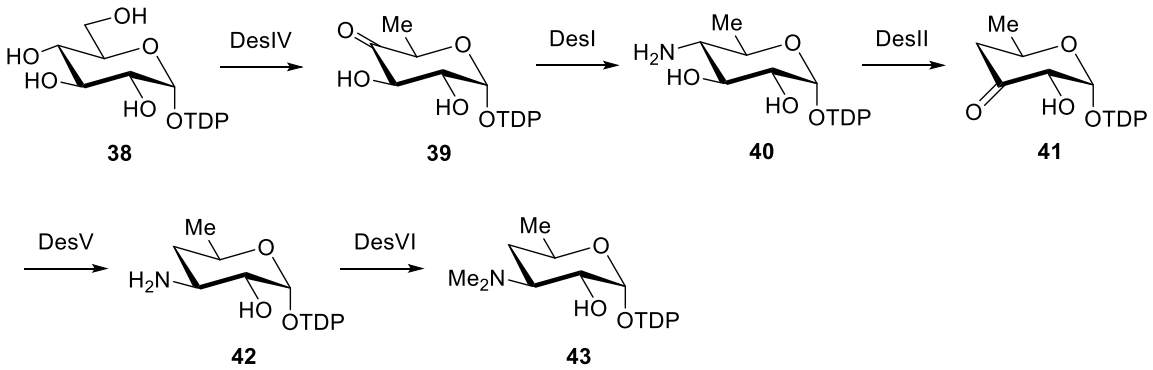


Figure 1-11. Biosynthetic pathway of TDP-desosamine (**43**).

Extensive biochemical characterization of DesII reaction strongly implicated the involvement of radical chemistry in the formation of C4 deoxy sugar **41**.[40] Interestingly, DesII was found to be capable of catalyzing both the redox neutral deamination of TDP-4-amino-4,6-dideoxy-D-glucose (**40**) to yield TDP-3-keto-4,6-dideoxy-D-glucose (**41**) and the oxidative dehydrogenation of TDP-D-quinovose (**44**) to yield TDP-3-keto-6-deoxy-D-glucose (**45**, see Figure 1-12).[41] Both reactions consume one equivalent of SAM per TDP sugar processed. While the deamination reaction requires an external source of reducing equivalents, the dehydrogenation reaction can regenerate the reduced $[4Fe–4S]^{1+}$ cluster in the absence of an external reductant during the catalytic cycle.

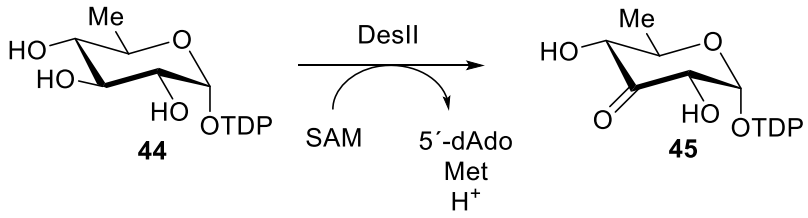


Figure 1-12. Dehydrogenation of TDP-D-quinovose (**44**) by DesII.

The initial hydrogen abstraction from the substrate occurs at C3. This has been confirmed by the observation that the deuterium of TDP-[3-^{2}H]-4-amino-4,6-dideoxy-D-glucose is incorporated into 5′-deoxyadenosine during catalysis by reconstituted DesII.[40] A radical intermediate that forms during the DesII-catalyzed dehydrogenation of TDP-D-quinovose has been characterized by EPR spectroscopy.[42] The doublet of doublets splitting pattern at a *g*-value of 2.0025 with hyperfine splitting constants of 33.6 G was detected. This EPR signal collapsed to a broad doublet when the C4-deuterated isotopologue (*i.e.*, TDP-D-[4-^{2}H]-quinovose) was incubated with DesII. Moreover, a reduction of the peak widths was observed when the reaction was conducted in D_2O instead of H_2O. These results

were consistent with the formation of an α-hydroxyalkyl radical centered at C3 with equivalent hyperfine couplings to the hydrogen atoms at both C2 and C4. The dihedral angle between the partially occupied *p*-orbital at C3 and the C–H bonds at C2 and C4 was estimated from coupling constants to be approximately 15°, which is not optimal for the elimination of substituents at C2 and C4. The measurement of solvent deuterium kinetic isotope effects using EPR spectroscopy and the pH titration of the dehydrogenation reaction provided additional mechanistic insights that deprotonation of the α-hydroxyalkyl radical takes place either prior to or in concert with electron transfer.[43]

Based on the aforementioned experimental evidence, a mechanism for the dehydrogenation of TDP-D-quinovose (**44**) by DesII has been proposed (Figure 1-13). This mechanism postulates an active site residue that can serve as a Lewis base and form a hydrogen-bonding interaction with the hydroxyl group at C3 of TDP-D-quinovose. Upon binding of SAM and the substrate, the reductive cleavage of SAM and hydrogen abstraction at C3 result in the generation of the α-hydroxyalkyl radical intermediate (**46**) observed by EPR. Deprotonation of the substrate radical by the active site base generates a ketyl radical (**47**), which may sufficiently lower the redox potential of substrate radical so as promote single electron transfer to the $[4Fe–4S]^{2+}$ cluster.[44,45] The electron transfer completes the dehydrogenation reaction with regeneration of the $[4Fe–4S]^{1+}$ cluster.

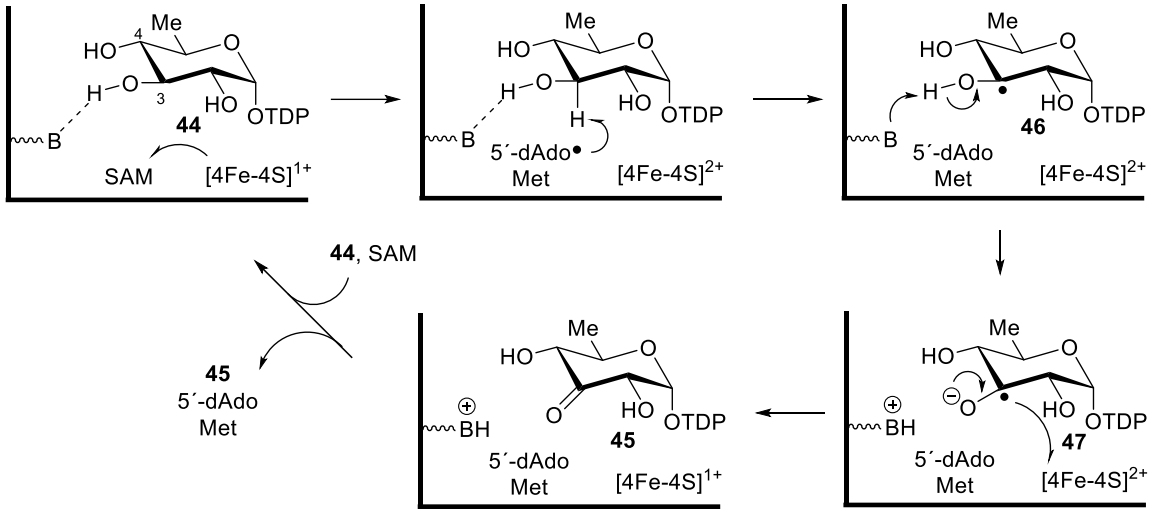


Figure 1-13. Proposed mechanism of dehydrogenation of TDP-D-quinovose (**44**) by DesII.

Although EPR characterization of a radical intermediate during the deamination reaction has not been successful yet, the current hypothesis proposes that the substrate radical formed during the deamination reaction adopts an alternative conformation compared to TDP-D-quinovose (Figure 1-14). It is proposed that TDP-4-amino-4,6-dideoxy-D-glucose (**40**) binds in the active site in such conformation that would allow a productive orbital overlap for elimination of amino group at C4. The exact mechanism for deamination after the formation of substrate radical (**48**) remains to be established; however, two possible mechanisms have been proposed: 1,2-migration and direct elimination. The first hypothesis involves migration of the amino group to C3 resulting in a carbinolamine radical intermediate (**49**), a process analogous to the radical deamination reaction catalyzed by B_{12}-dependent ethanolamine ammonia lyase (EAL) (Pathway A).[46,47] Reduction of the resulting carbinolamine radical and elimination of ammonia generate the corresponding C3 ketosugar (**41**). The alternative hypothesis is a direct elimination

mechanism, in which deprotonation of the substrate radical is followed by the direct elimination of amino group at C4. However, deprotonation would seem to promote dehydrogenation rather than deamination, as is the case with TDP-D-quinovose. This conflict can be resolved by catalytic model in which the active site general base does not deprotonate the substrate or substrate radical but rather serves to facilitate a proton transfer. In this case, proton transfer from the C3 hydroxyl group to the C4 amino group proceeds in concert with the elimination of ammonia (Pathway B, **50**). A computational study of EAL-catalyzed transformation also suggests that the energy barrier for direct loss of an ammonium cation is comparable to that for migration of a protonated amine group.[47] Regardless of which mechanism is operating, the deamination reaction requires an external electron source to reduce the product radical and regenerate the $[4Fe–4S]^{1+}$ cluster during each catalytic cycle.

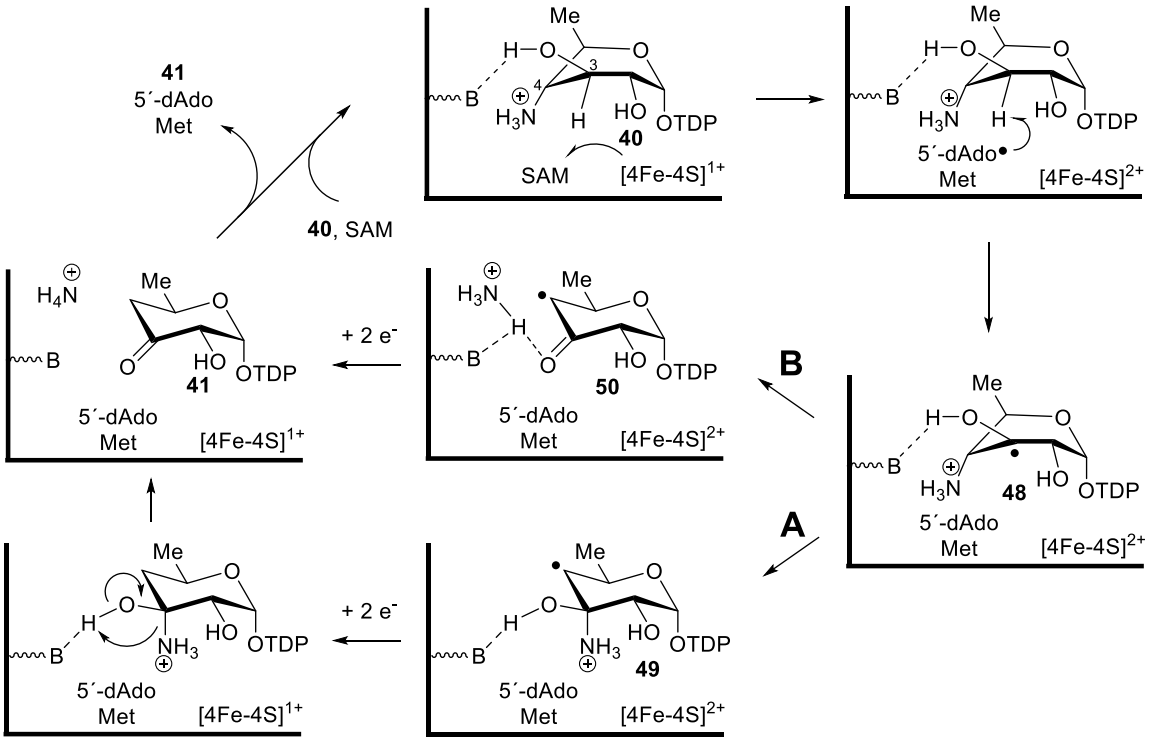


Figure 1-14. Proposed mechanism of deamination of TDP-4-amino-4,6-dideoxy-D-glucose (**40**) by DesII.

1.2 BIOSYNTHESIS OF *C*-NUCLEOSIDES

1.2.1 *C*-Nucleosides

Many fundamental biological activities such as the storage of genetic information, control of signaling pathways, and transfer of intracellular energy are mediated by nucleosides and nucleotides. A number of modified nucleosides from secondary metabolites exhibit a variety of biological properties including antibacterial, antiviral, and antitumor activities.[48,49] Among them, *C*-nucleosides in which the nucleobase is attached to a sugar moiety via *C*-glycosidic bond have garnered specific attention due to their therapeutic potential.[50] Therefore, studies of the biosynthetic pathways of these nucleosides and enzyme mechanisms involved in their formation are of interest.

Many *C*-nucleosides have been identified. For example, 8-aza-9-deazaadenosine (**51**), commonly known as formycin A, is an antibiotic isolated from *Nocardia interforma* and *Streptomyces kaniharaensis* SF-557 (see Figure 1-15).[51,52] Formycin A is a structural isomer of adenosine (**53**) having a unique pyrazolopyrimidine nucleobase linked to ribose via a *C*-glycosidic bond in contrast to adenosine where an imidazolopyrimidine nucleobase is linked to ribose via an *N*-glycosidic bond. In a similar way, formycin B (**52**) is the corresponding *C*-nucleoside analog of inosine (**54**).[53]

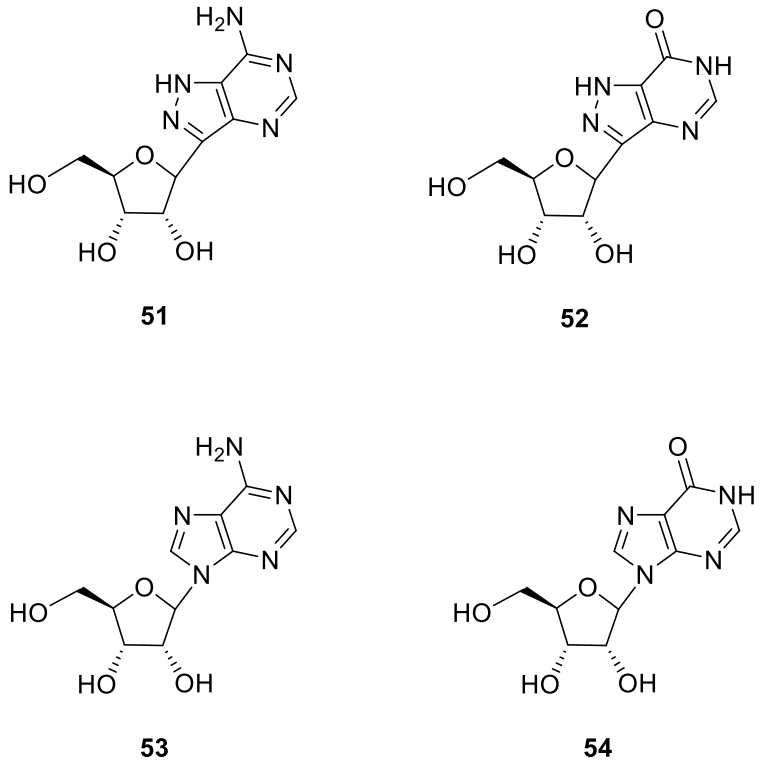


Figure 1-15: Structures of formycin A (**51**), formycin B (**52**), adenosine (**53**), and inosine (**54**).

There are several related *C*-nucleoside antibiotics, which include pyrazomycin (**55**), showdomycin (**56**), minimycin (**57**) and pseudouridine (**58**) (see Figure 1-16). Pyrazomycin (or pyrazofurin, **55**), produced by *Streptomyces candidus*, is a pyrazole nucleoside that exhibits antiviral and antitumor activities.[54] Showdomycin (**56**) isolated from *Streptomyces showdoensis*[55] and minimycin (or oxazinomycin, **57**) isolated from *Streptomyces hygroscopicus*[56] contain a maleimide ring and an oxazinedione ring, respectively. Pseudouridine (**58**) is a post-transcriptionally modified derivative of uridine found in RNA.[57] The uracil moiety in pseudouridine is also attached to the ribose moiety via a *C*-glycosidic bond.

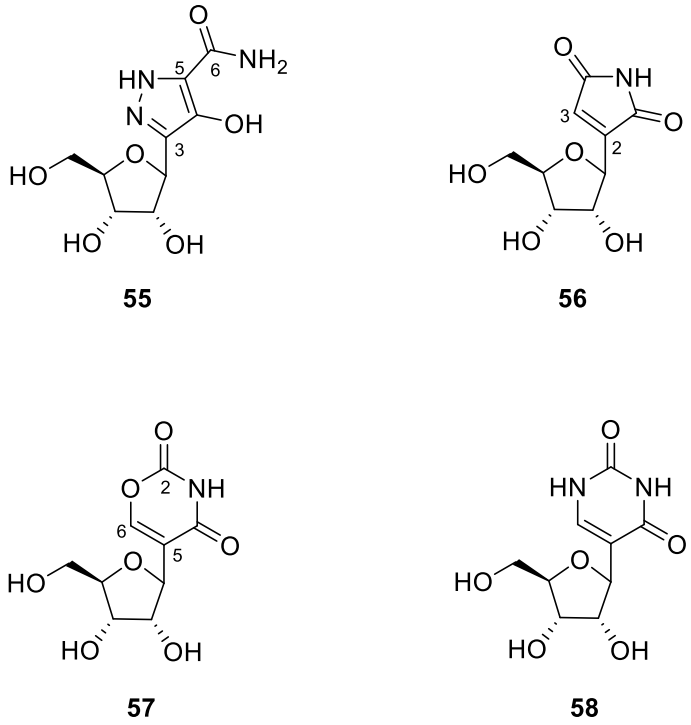


Figure 1-16: Structures of representative *C*-nucleosides: pyrazomycin (**55**), showdomycin (**56**), minimycin (**57**) and pseudouridine (**58**).

1.2.2 Biosynthesis of Pseudouridine

While the biogenesis of *C*-nucleosides in nature remains an active area of investigation, pseudouridine (**58**) represents one example where the biosynthetic pathway has been established. Pseudouridine is derived from uridine in RNA (**59**) upon isomerization by pseudouridine synthase as shown in Figure 1-17.[58] The proposed mechanism involves an essential aspartate residue that operates as a general base.[59] Elimination of the uracil moiety may be either stepwise followed by the deprotonation of oxocarbenium intermediate (Pathway A), or concerted with deprotonation at C2′ (Pathway B), which generates the key glycal intermediate **60**. The pyrimidine ring can then rotate in the active site prior to addition back to the ribose to form the *C*-glycoside.

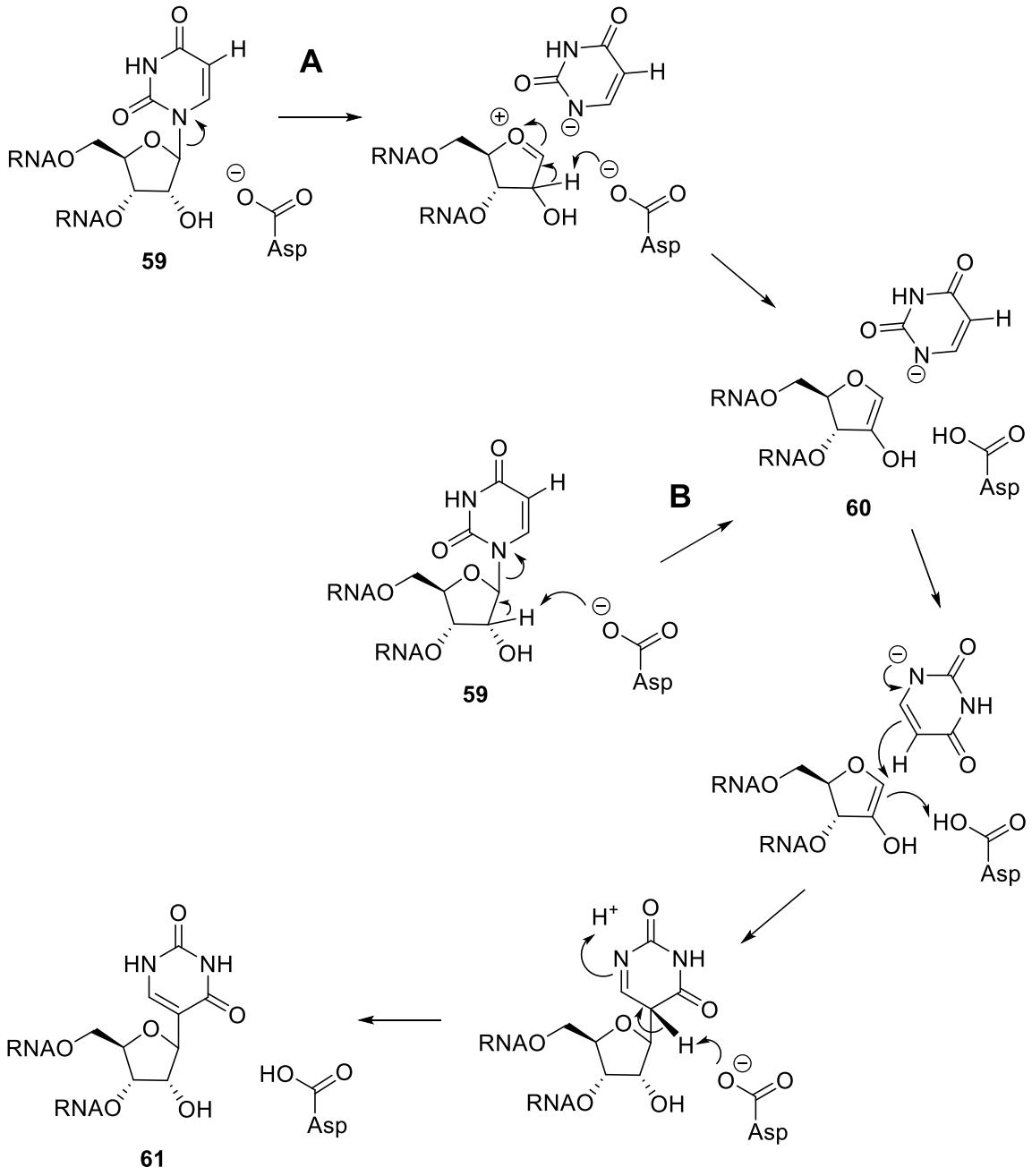
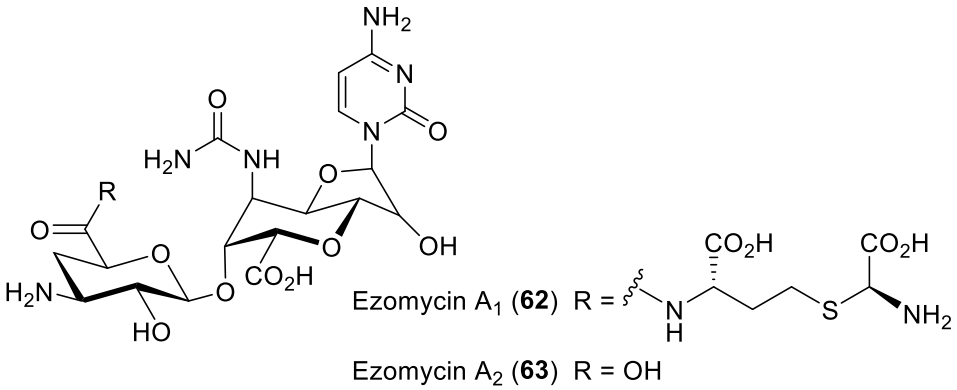
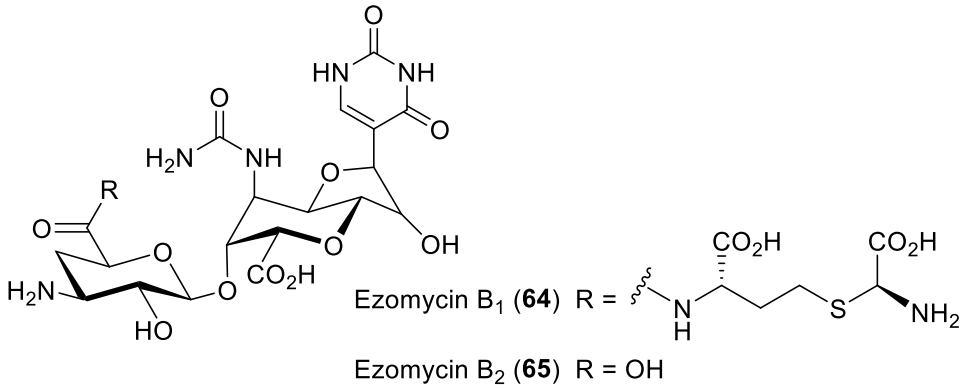
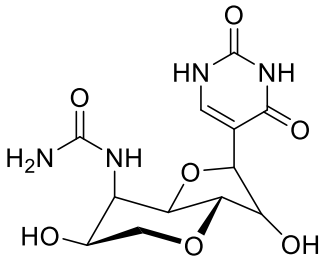


Figure 1-17: Mechanistic proposals for the biosynthesis of pseudouridine (**58**) from uridine.

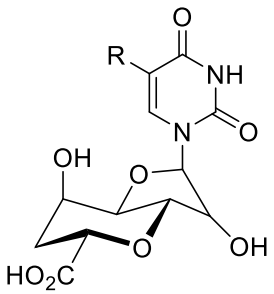
Ezomycins produced by a *Streptomyces* species are bicyclic uronic acid nucleosides that exhibit antifungal activity.[60] Whereas ezomycin A_1 (**62**) and A_2 (**63**) are cytosine *N*-nucleosides that share the octosyl acid core scaffold, ezomycin B_1 (**64**) and B_2 (**65**) are *C*-nucleosides containing a pseudouridine moiety instead of cytosine (Figure 1-18).[61,62] Structurally related, malayamycin A (**66**) isolated from *Streptomyces malaysiensis* is a bicyclic *C*-nucleoside with strong antifungal activity.[63] Biosynthesis of the octosyl acid core scaffold (**73**) was proposed to be derived from phosphoenolpyruvate (**70**, PEP) and a uridine derivative based on the isotope labeling studies.[64] *In vitro* characterization of the radical SAM enzyme PolH and the phosphatase PolJ provided evidence for a biosynthetic pathway of octosyl acid that begins with uridine 5′-monophosphate (**69**, UMP, see Figure 1-19).[65] As of yet, no biosynthetic studies have been conducted to elucidate the biotransformations leading to the ezomycins or malayamycin A. However, it is conceivable that an isomerase similar to pseudouridine synthase may operate to construct the *C*-glycosidic linkage, given the fact that both *C*-nucleosides and the structurally analogous *N*-nucleosides have been isolated together from the same culture filtrates.







Malayamycin A (**66**)



Octosyl acid A (**67**) R = CO_2H
Octosyl acid B (**68**) R = CH_2OH

Figure 1-18. Structures of ezomycins (**62**, **63**, **64**, **65**), malayamycin A (**66**) and octosyl acids (**67**, **68**).

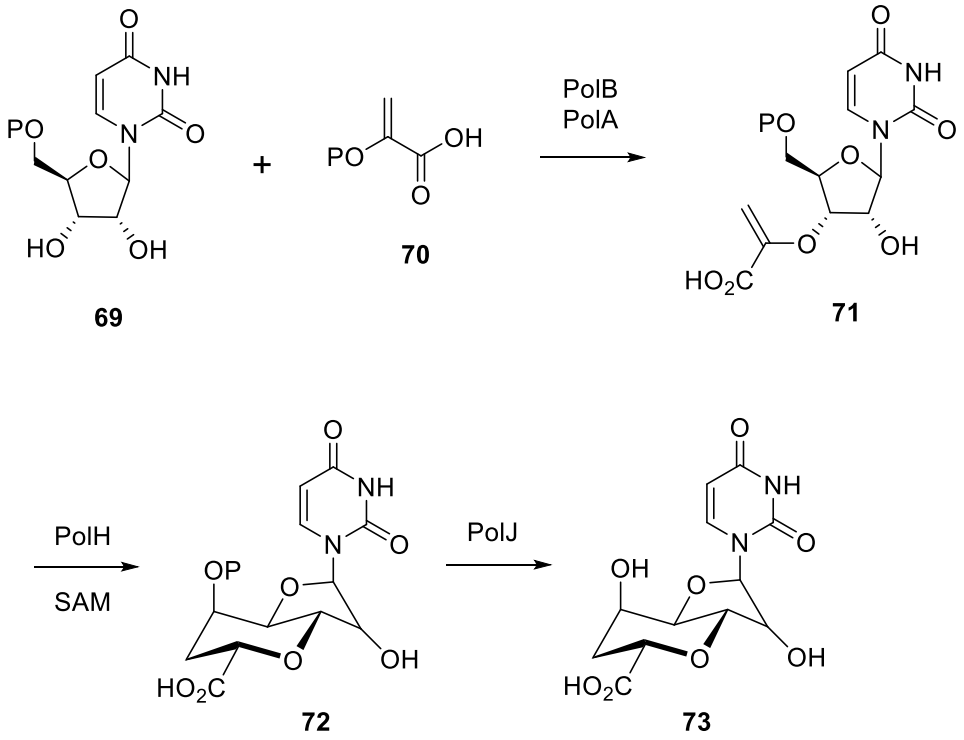


Figure 1-19. Proposed pathway for the octosyl acid core scaffold (**73**).

1.2.3 Biosynthesis of Formycin A and Related *C*-Nucleosides

The biosynthetic pathway leading to the formation of the *C*-glycosidic bond found in the formycins (**51**, **52**), pyrazomycin (**55**), showdomycin (**56**), and minimycin (**57**) remains elusive. Early studies of formycin A biosynthesis have identified a sequence of chemical transformation from formycin B 5′-phosphate (**77**) to formycin A (**51**) *via* formycin A 5′-phosphate (**78**) that are proposed to take place *in vivo* (see Figure 1-20).[66] Isotope tracing experiments in *N. interforma* have shown that the ribosyl moiety of formycin A is derived directly from ribose likely via an intermediary produced from phosphoribosyl pyrophosphate (**74**, PRPP).[67] The *ε*-amino group of L-lysine (**76**) was found

to be the source of the N3, N7, and N8 nitrogen centers of formycin A.[68,69] Importantly, the C6, C5, C4, and C9 carbons of formycin A originate from the C1, C2, C3, and C4 carbons of L-glutamate (**75**), implying that L-glutamate is asymmetrically incorporated into formycin A as an intact precursor (Figure 1-20).[69,70] The biosynthetic origins of formycin A differ from those for adenosine (**53**) where the C4, C5 and N7 centers are from L-glycine (**79**, see Figure 1-21).[71] The other two nitrogens on the adenine ring of **53**, namely N3 and N9, are derived from L-glutamine (**80**), while bicarbonate is the source of the C6 carbon, and the C2 and C8 centers originate from N^{10}-formyl-tetrahydrofolate (N^{10}-formyl-THF). Thus, the biosynthetic pathway for formycin A (**51**) appears to be distinct from that of adenosine (**53**) despite their significant structural resemblance.

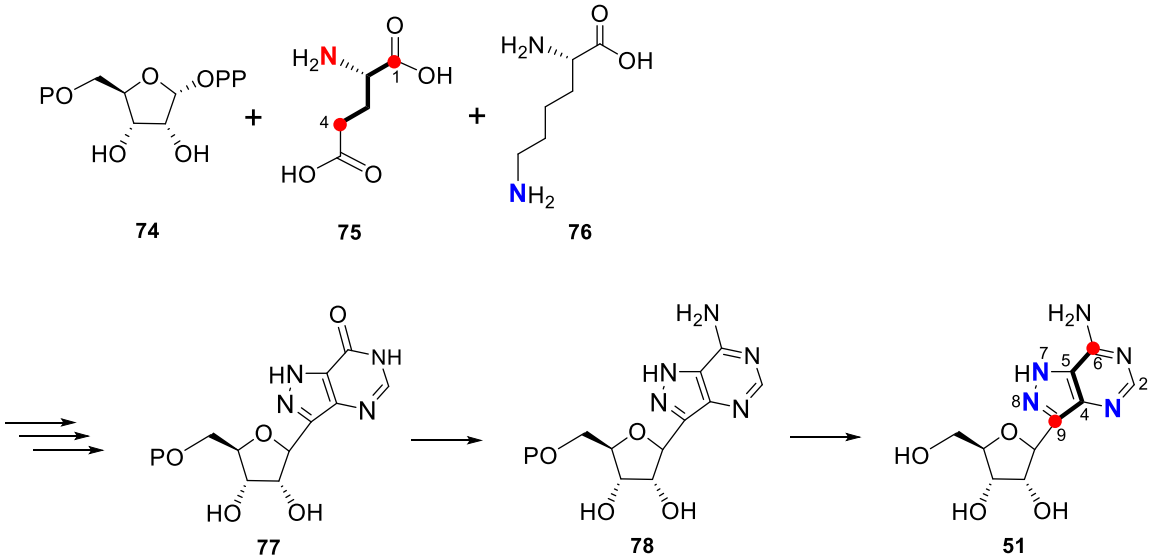


Figure 1-20. Biosynthetic origins of formycin A (**51**).

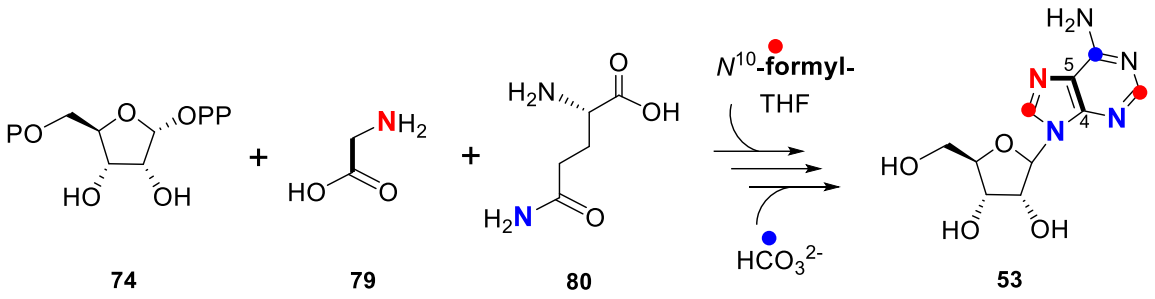


Figure 1-21. Biosynthetic origins of adenosine (**53**).

L-glutamate (**75**) appears to be a common precursor for the formycin class of nucleoside antibiotics, given similar isotope incorporation patterns. The four adjacent carbons of formycin A are directly derived from C1 to C4 of L-glutamate as are all of the carbons of the pyrazole ring and amide carbon of pyrazomycin (**55**).[70] Furthermore, the precursor of the maleimide ring in showdomycin (**56**) and that of the oxazinedione ring in minimycin (**57**) have also been shown to be L-glutamate by feeding experiments. The maleimide moiety of showdomycin is constructed from the C2 to C5 four-carbon unit and nitrogen from L-glutamate.[72] Likewise, the three carbons from C3 to C5 of L-glutamate are asymmetrically incorporated into C6 to C4 of the oxazine ring of minimycin.[73] A common intermediate for *C*-nucleosides is thus hypothesized and believed to be formed via coupling of PRPP (**74**) and L-glutamate (**75**) followed by the loss of either the C1 or C5 carbon of L-glutamate upon decarboxylation.

The biosynthesis of *C*-nucleosides seems to be different from that of other naturally occurring aryl *C*-glycosides where the sugar moiety is attached to the aromatic ring by a C-C bond.[74] Although the mechanistic features of enzymatic *C*-glycosyl transfers are not completely understood in comparison to those of *O*-glycosyltransfers,[75] several *C*-glycosyltransferases have been reported and characterized,[76-78] which include IroB from *E. coli*. IroB catalyzes three consecutive glucosyl transfers from uridine 5′-diphosphoglucose (**82**, UDP-Glc) to enterobactin (**81**, see Figure 1-22).[79] The *C*-glycosylation by IroB has been proposed to proceed via a Friedel-Crafts like reaction mechanism as shown in Figure 1-23.[80] The C-C bond is formed via nucleophilic attack to the anomeric carbon of an NDP-sugar donor and is facilitated by electron-donating groups on the aromatic ring. A study of the catalytic promiscuity of *C*-glycosyltransferase indicated that the regiospecificity of *C*-glycosylation is determined by the chemical properties and the locations of substituents on the aromatic ring.[81] The positioning of the sugar acceptor by active site residues also plays a significant role.[82,83] In contrast, the assembly of *C*-nucleosides is less likely to follow the aromatic substitution mechanism, because their heterocyclic nucleobases lack electron-donating substituents as directing groups.

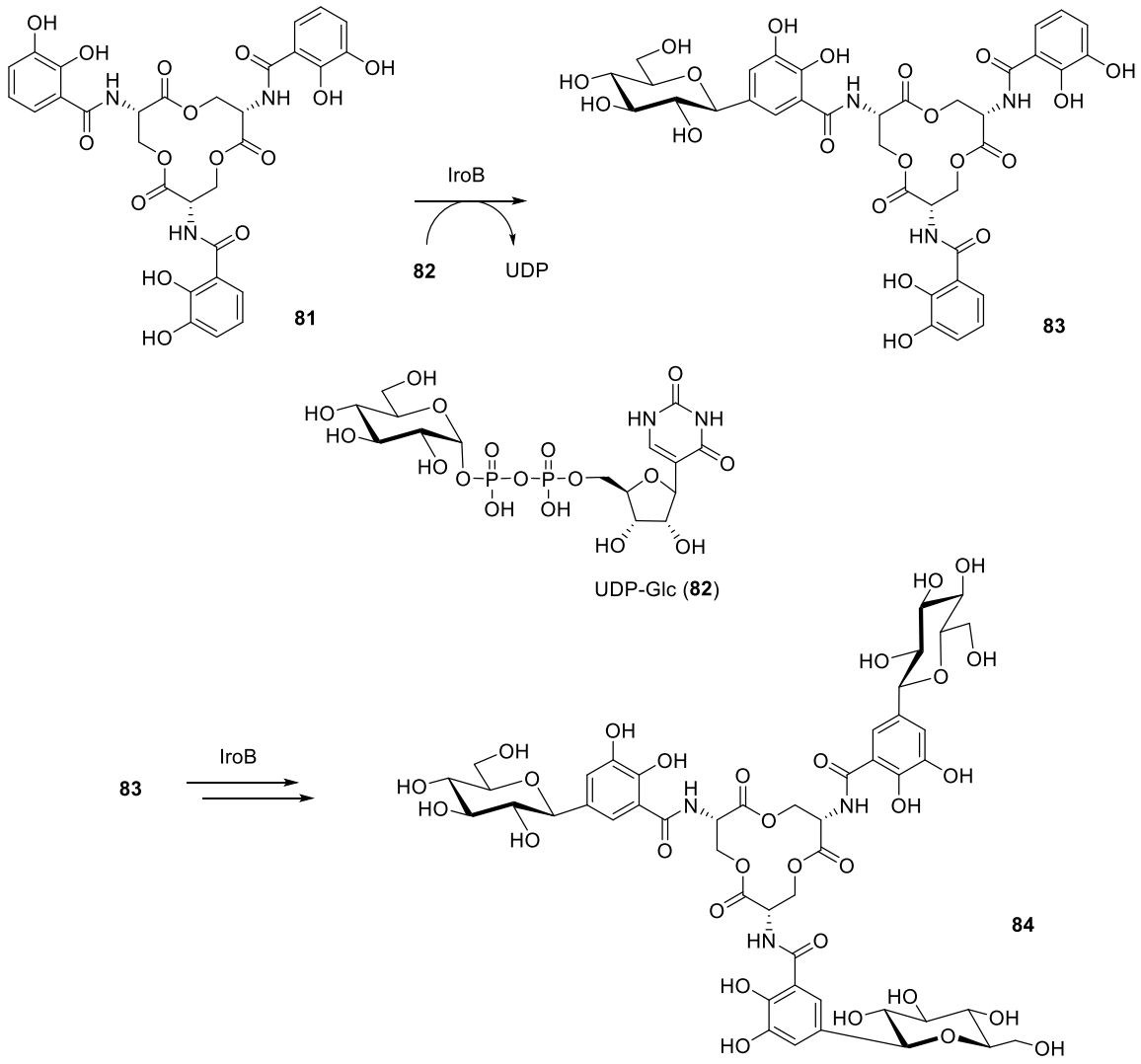


Figure 1-22. IroB-catalyzed glucosylations of enterobactin (**81**).

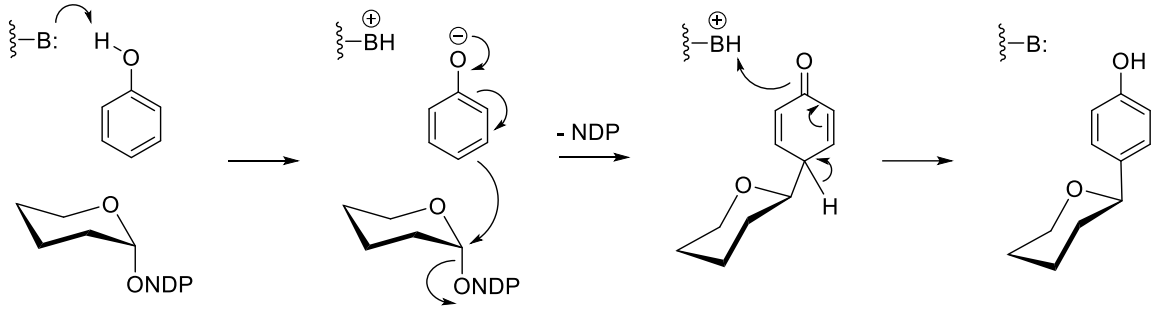


Figure 1-23. Proposed mechanism of *C*-glycosylation in the biosynthesis of *C*-glycosides.

1.3 SUMMARY AND THESIS STATEMENT

The radical SAM enzyme DesII can operate either as a lyase or a dehydrogenase depending on the substrate provided. The dual functionality of DesII raises the question of how the fate of the radical intermediate is governed during the catalytic cycle. It is hypothesized that substrate geometry in the active site plays an important role in controlling the partitioning of the radical intermediate into different reaction pathways. Specifically, the mechanistic difference between deamination and dehydrogenation reactions may be caused in large part by the degree of orbital overlap in the bound radical intermediate. In Chapter 2, mechanistic investigation of DesII reactions with several substrate analogs is described. It was shown that all analogs with altered stereochemistry can be processed by DesII, despite being poor substrates. Characterization of each product demonstrated that the reaction partitions into alternative pathways upon changing the stereochemistry of the ring substituents of the substrate. This study provides experimental evidence that the stereochemical configuration of the substrate radical significantly contributes to the reaction outcomes during the catalytic cycle of DesII.

Formycin A, isolated from *N. interforma* and *S. kaniharaensis*, is a *C*-nucleoside analog of adenosine. The biosynthesis of nucleosides is of particular interest because of their unique structural features such as the pyrazolopyrimidine moiety with a *C*-glycosidic linkage in formycin A. Chapter 3 focuses on the identification and characterization of the biosynthetic gene cluster for formycin A. In order to investigate the biosynthesis of formycin A, the genome of *S. kaniharaensis* was sequenced. Further analysis led to a plausible route for the assembly of the pyrazolopyrimidine moiety in formycin A. The biosynthetic gene cluster was later identified by construction and screening of a cosmid library. Bioinformatics analysis permitted assignment of putative functions of each gene, from which a biosynthetic pathway for formycin A could be proposed. Several of the enzymes involved in the biosynthesis of formycin A have been cloned, expressed, and assayed for activity. The ultimate goal of this research is to fully elucidate the entire biosynthetic pathway of formycin A, and investigate the detailed reaction mechanisms of the enzymes involved.

Chapter 2: Mechanistic Investigation of the Radical SAM Enzyme DesII Using Substrate Analogs with Altered Stereochemistry

2.1 INTRODUCTION

DesII is the radical SAM enzyme from *S. venezuelae* that catalyzes the deamination of TDP-4-amino-4,6-dideoxy-D-glucose (**40**) to produce TDP-4,6-dideoxy-3-keto-D-glucose (**41**) (see Figure 2-1).[40] The deamination is the key reaction in the biosynthesis of TDP-desosamine (**43**), which is an indispensable structural component required for the biological activity of many macrolide antibiotics.[35] The radical-mediated deamination reaction catalyzed by DesII represents one of the regioselective deoxygenation reactions of a glucose derivative to achieve structural diversity. However, when DesII is presented with the nonbiosynthetic substrate TDP-D-quinovose (**44**), in which the C4 amino group is replaced by a hydroxyl group, it catalyzes oxidation of the C3 hydroxyl group to generate TDP-3-keto-6-deoxy-D-glucose (**45**).[41] The oxidative dehydrogenation of **44** is analogous to the oxidation of 2-deoxy-*scyllo*-inosamine (**7**) catalyzed by the radical SAM enzyme BtrN in the biosynthesis of butirosin (see Figure 1-2 in Chapter 1.1.2),[9] which is in stark contrast with the redox-neutral deamination of **40**. Moreover, it is unexpected that DesII does not catalyze the dehydration of **44** given the precedence set by the dehydration of 1,2-diols by cobalamin-dependent dioldehydratases.[84] Furthermore, given the fact that nonenzymatic dehydration of α,β-dihydroxyalkyl radicals was observed in solution also suggests that the dehydration of radical adducts of **44** should be rapid.[85-87] Therefore, study

---

The part of this chapter has been published, and I was the primary author (Ko, Y.; Ruszczycky, M. W.; Choi, S.-H.; Liu, H.-w. Mechanistic studies of the radical *S*-adenosylmethionine enzyme DesII with TDP-D-fucose. *Angew. Chem. Int. Ed.* **2015**, *54*, 860–863).

of the mechanistic details underlying the dual activity of DesII may shed light on the subtleties of radical control during the catalytic cycle.

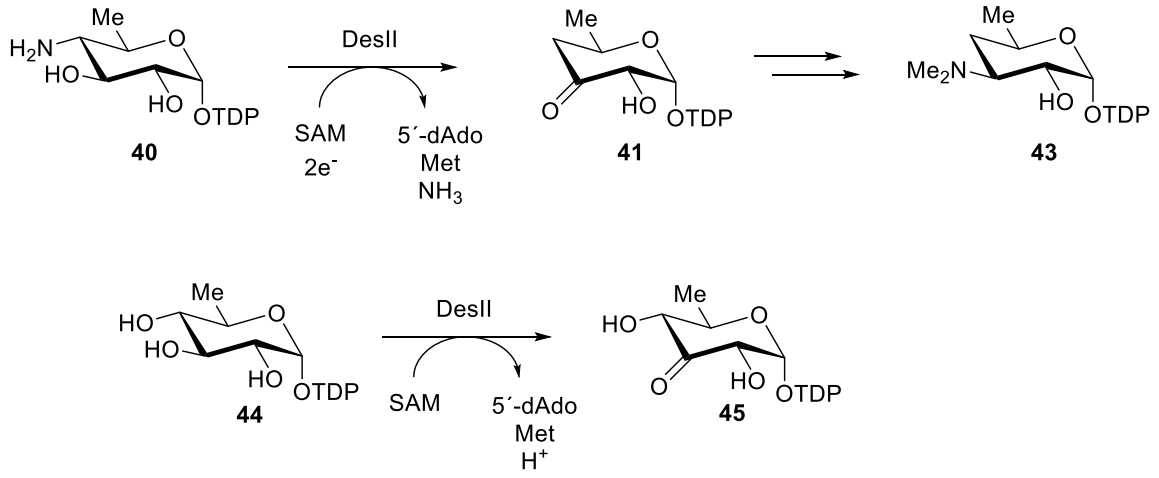


Figure 2-1. Reactions catalyzed by DesII.

The radical intermediate formed during the dehydrogenation of **44** was identified by EPR spectroscopy to be an *α*-hydroxyalkyl radical centered at C3 (**46**, see Figure 1-13).[42] The C–O bonds at both C2 and C4 are essentially orthogonal to the partially filled *p*-orbital at C3, which should impede elimination of the hydroxyl groups. On the contrary, a different binding configuration of the substrate radical from **40** is proposed in which the C4 amino group is better overlapped with the *p*-orbital at C3 allowing the deamination reaction to take place.[88] It is hypothesized that the degree of orbital interaction, which depends on the binding configuration of the parent compound in the active site, plays a key role in determining whether the enzyme operates as a lyase or a dehydrogenase (see the discussion in Chapter 1.3). When sufficient hyperconjugation between the C4 substituent

and the C3 *p*-orbital is provided, the elimination reaction can occur to afford deamination/dehydration products. If the binding configuration does not allow a productive orbital overlap so that elimination is impeded, competing processes such as deprotonation of the C3 hydroxyl group by an active site base may become significant.[43] In this case, the electron transfer to the $[4Fe–4S]^{2+}$ cluster would be facilitated by the more strongly reducing ketyl radical (**47**) and the reaction should then proceed with elimination of the C4 substituent.

This working hypothesis implies that perturbation of the hyperconjugation between the singly occupied *p*-orbital at C3 and the C4 C–N/O bond of the radical intermediate can modulate partitioning between the elimination and oxidation pathways.[88] In order to investigate this hypothesis, the reaction of DesII with several diastereomeric substrate analogs were examined and the results are described in this chapter. The substrate analogs employed in this study include TDP-4-amino-4-deoxy-D-fucose (**85**), TDP-D-fucose (**86**), and TDP-6-deoxy-D-mannose (**87**, see Figure 2-2). Compounds **85** and **86** represent the C4 epimers of TDP-4-amino-4,6-dideoxy-D-glucose (**40**) and TDP-D-quinovose (**44**), respectively, and their reactions with DesII could provide mechanistic insight regarding the influence of the stereochemistry at C4. Furthermore, the fate of **87** is of interest, because the inverted configuration at C2 may facilitate elimination of the C2 substituent. Reported herein is the experimental evaluation of the DesII reaction with **85**, **86**, and **87**. The study using **86** as a potential substrate has been published, and the results and discussion are directly adapted from that publication.[89]

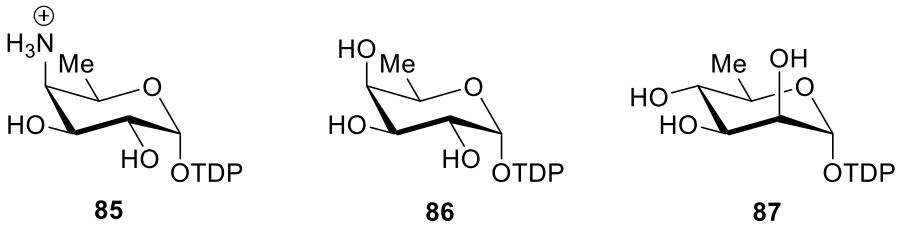


Figure 2-2. Structures of substrate analogs employed in this study.

2.2 Experimental Procedures

2.2.1 General

All reactions involving DesII were performed in a Coy anaerobic chamber (glovebox) under an atmosphere of approximately 97.5% N_2 and 2.5% H_2 with less than 1 ppm O_2. All chemicals were obtained from commercial sources and used without further purification unless specified otherwise. Except for those prepared with D_2O (see Chapter 2.2.4.2), all solvents were deaerated by bubbling with nitrogen gas before being transferred into the glovebox. Once in the glovebox, the solvents were stirred open to the anaerobic atmosphere overnight to allow equilibration. All electrospray ionization mass spectrometry (ESI-MS) was performed at the Mass Spectrometry Facility of the University of Texas at Austin using an Agilent 6530 Q-TOF MS. All NMR spectroscopy was performed at the NMR Facility of the University of Texas at Austin.

2.2.2 Enzymes

2.2.2.1 DesII

DesII from *S. venezuelae* containing a *C*-terminal His_6 tag was heterologously overexpressed in BL21 Star *E. coli* using a pET24b(+) expression plasmid and purified aerobically according to previously published methods.[40] The purified enzyme in 50 mM potassium phosphate buffer (pH 7.5) containing 15% glycerol was flash frozen and stored at –80 °C in approximately 0.5 mL aliquots containing roughly 2.5 mg protein each.

Reconstitution also followed previously published methods.[41,42] Prior to reconstitution, approximately 3–5 mg of enzyme was thawed and dialyzed overnight versus 25 mM 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS) buffer (pH 8.0)

containing 10% glycerol under a N_2 atmosphere at 4 °C. The dialysis system was then transferred to the glovebox. There, the dialyzed enzyme was diluted to a volume of ca. 2 mL and gently stirred at ca. 10 °C in an open conical vial in order to equilibrate with the anaerobic atmosphere. The solution was then made 5 mM in dithiothreitol (DTT) and allowed to stir for 15 min before adding SAM to 1.5 mM and allowing the reaction to stir for another 5 min. Next, 100 µL of a 40 mM solution of $Fe(NH_4)_2(SO_4)_2$ in anaerobic H_2O was slowly added over 10 min. This was followed by the slow addition of 100 µL of a 40 mM solution of Na_2S in anaerobic 25 mM EPPS buffer (pH 8.0) over 20 min. Once all of the Na_2S had been added, the solution was gently stirred at ca. 10 °C for 5 h. The reconstituted enzyme was then loaded onto a 50 mL Sephadex G-25 column prequilibrated with anaerobic buffer containing 25 mM EPPS (pH 8.0) and 1 mM DTT, and the column was eluted with the same buffer. Brown colored fractions were collected, pooled, concentrated to approximately 30 µM enzyme and stored in the glovebox at ca. 10 °C until use.

2.2.2.2 FdtA

The ketoisomerase FdtA containing an *N*-terminal His_6 tag was heterologously overexpressed and purified as previously described.[90] The *fdtA* gene from *Aneurinibacillus thermoaerophilus* was PCR amplified and cloned into the pET28b(+) vector at the NdeI and EcoRI restriction sites. The primers used were: 5′-GGGTAGGGCATATGGAAAA-TAAAG-3′ (forward, with the NdeI site underlined) and 5′-CCCGAATTCTTATCCTTC-TTTTTC-3′ (reverse, with the EcoRI site underlined). The resulting plasmid was used to transform *E. coli* BL21 Star (DE3) cells.

An overnight culture of *E. coli* transformed with the *fdtA*/pET28b(+) plasmid was then grown at 37 °C in LB medium containing 50 µg/mL kanamycin and used to inoculate

six 1 L of the same medium in a 100-fold dilution. The 1 L cultures were incubated at 37 °C until the OD_{600} reached 0.6, at which point protein expression was induced by the addition of isopropylthiogalactoside (IPTG) to a final concentration of 0.2 mM, and the temperature was decreased to 18 °C. Following overnight incubation at 18 °C, the cells were harvested by centrifugation at 4500 × g for 8 min, resuspended in 80 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and disrupted by sonication. Cell debris was removed by centrifugation at 20000 × g for 30 min, and the supernatant was mixed with 8 mL of nickel-nitrilotriacetic acid (Ni-NTA) resin for 10 min under slow agitation at 4 °C. The slurry was transferred to an empty column and washed with 100 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 20 mM imidazole. The protein was eluted with 50 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 250 mM imidazole and dialyzed twice against 1 L of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 20% glycerol. The dialyzed protein was then flash frozen with liquid N_2 and stored at –80 °C until use.

2.2.2.3 RfbB, DesV, and Tyl1a

The TDP-glucose 4,6-dehydratase RfbB from *Salmonella typhi*,[91] the transaminase DesV from *S. venezuelae*,[36] and the 3,4-ketoisomerase Tyl1a from *Streptomyces fradiae*[92] were overexpressed and purified as previously described.

2.2.3 Compounds

2.2.3.1 SAM and [5',5'-$^2H_2$]SAM

SAM was prepared from adenosine triphosphate (ATP) and L-methionine using the SAM synthetase MetK and isolated using cation exchange batch phase extraction as previously described with minor variations.[93,94] The dideuterated isotopolog *S*-[5′,5′-

$^{2}H_2$]adenosyl-L-methionine ([5′,5′-$^{2}H_2$]SAM) was prepared analogously to natural abundance SAM using [5′,5′-$^{2}H_2$]adenosine triphosphate ([5′,5′-$^{2}H_2$]ATP) and L-methionine. The HRMS (ESI, positive ion mode) calculated for the product $C_{15}H_{21}D_2N_6O_5S^+$ $[M]^+$ was 401.1569 *m*/*z*, versus the observed 401.21 *m*/*z*. [5′,5′-$^{2}H_2$]ATP was synthesized from adenosine (**53**) according to previously described procedures by Dr. Sei Hyun Choi.[95]

2.2.3.2 DesII substrate analogs: TDP-4-amino-4-deoxy-D-fucose, TDP-D-fucose, and TDP-6-deoxy-D-mannose

The deuterated substrate analogs, TDP-[3-^{2}H]-4-amino-4-deoxy-D-fucose (**85D**) and TDP-[3-^{2}H]-6-deoxy-D-mannose (**87D**) were chemically synthesized by Geng-Min Lin (*manuscript in preparation*). TDP-D-fucose (**86**) was prepared and purified as previously described (see Figure 2-3).[40]

2.2.3.3 TDP-6-deoxy-D-gulose

TDP-6-deoxy-D-gulose standard (**98**) was prepared from TDP-D-glucose (**38**) as shown in Figure 2-3. A 10 mL solution containing 30 mM TDP-D-glucose (**38**), 0.3 mM NAD, 50 mM Tris (pH 7.5), and ca. 1 mg RfbB was incubated at 37 °C for 5 h. Next, ca. 3 mg FdtA was added to the reaction mixture, and the solution was incubated at room temperature overnight. After centrifugation and filtration using a Millipore YM-10 ultrafiltration membrane, 20 mM $NaBH_4$ was added and the reaction mixture was incubated for 1 h on ice. The FdtA reaction equilibrium favors TDP-6-deoxy-4-keto-D-glucose (**39**); therefore, reduction by $NaBH_4$ generated a mixture of TDP-D-quinovose (**44**), TDP-D-fucose (**86**) and TDP-6-deoxy-D-gulose (**98**). The desired product (**98**) was purified by HPLC using a Dionex CarboPac PA1 semiprep column. A two-solvent system of A, H_2O, and B, 1.5 M ammonium acetate in H_2O (pH 7.0), was used with a flow rate of 3 mL/min.

The column was eluted with a two-phase linear gradient: 2.5 to 15% B in 20 min, 15 to 20% B in 20 min and return to 2.5% B in 5 min. Although TDP-D-fucose (**86**) could be removed in this manner, complete separation of TDP-D-quinovose (**44**) from TDP-6-deoxy-D-gulose (**98**) was not possible, because the elution times were very similar at ca. 25.5 min. Nevertheless, two consecutive HPLC purifications produced a 2:1 mixture of TDP-6-deoxy-D-gulose to TDP-D-quinovose. NMR characterization of the TDP-D-quinovose was consistent with previous reports permitting deconvolution of spectroscopic data for TDP-6-deoxy-D-gulose.[96] ^{1}H NMR and ^{31}P NMR spectra are shown in Figure 2-4 and Figure 2-5, respectively. ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 7.58 (s, 1H, 6-H), 6.19 (dd, 1H, J = 7.0, 7.0 Hz, 1′-H), 5.38 (dd, 1H, J = 3.7, 6.8 Hz, 1″-H), 4.45–4.49 (m, convoluted with HOD peak, 1H, 3′-H), 4.29 (q, 1H, J = 6.6 Hz, 5″-H), 4.00–4.06 (m, 3H, 4′-H, 5′-H), 3.82 (dd, 1H, J = 3.7, 3.7 Hz, 3″-H), 3.76 (ddd, 1H, J = 3.7, 3.7, 3.7 Hz, 2″-H), 3.59 (d, 1H, J = 3.7 Hz, 4″-H), 2.18–2.26 (m, 2H, 2′-H), 1.77 (s, 3H, CH_3), 1.07 (d, 3H, J = 6.6 Hz, 6″-H); ^{31}P NMR (D_2O, 243 MHz) δ (ppm) –11.31 (d, J = 20.0 Hz), –13.03 (d, J = 20.0 Hz). HRMS (ESI, negative ion mode) calculated for $C_{16}H_{25}O_{15}N_2P_2^-$ $[M-H]^-$: 547.0736 *m*/*z*, found: 547.0723 *m*/*z*.

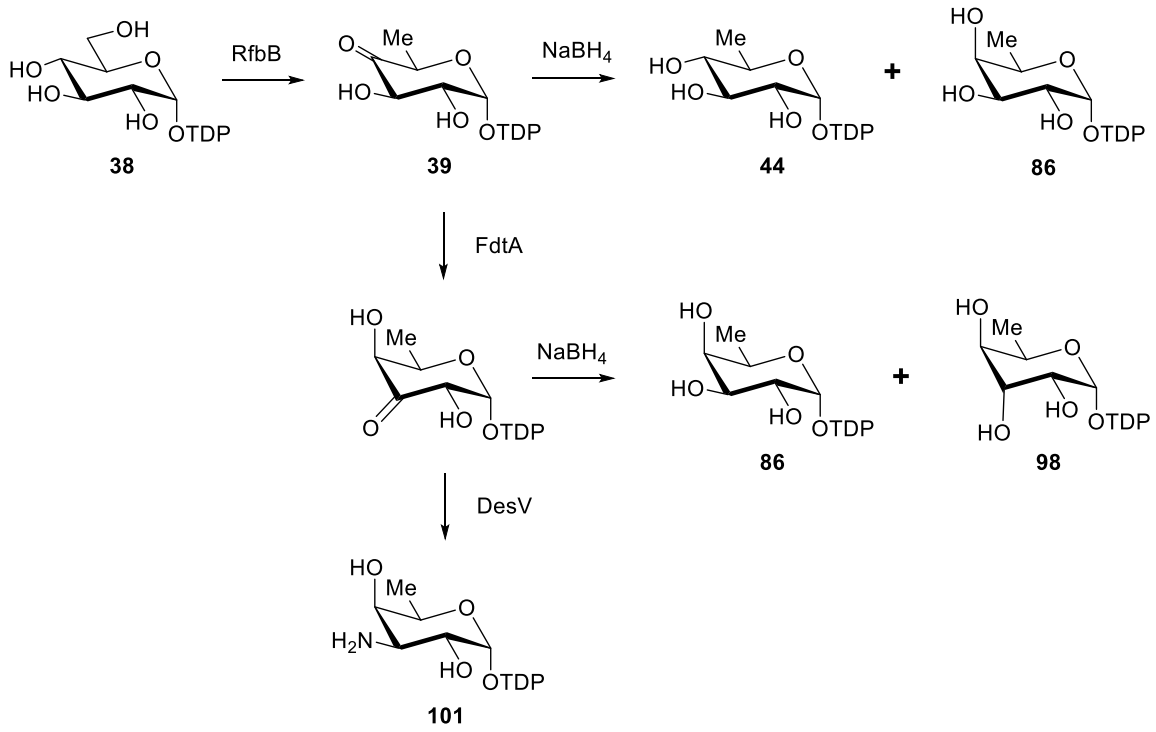


Figure 2-3. Preparation of TDP-6-deoxy-D-gulose (**98**) and TDP-3-amino-3-deoxy-D-fucose (**101**).

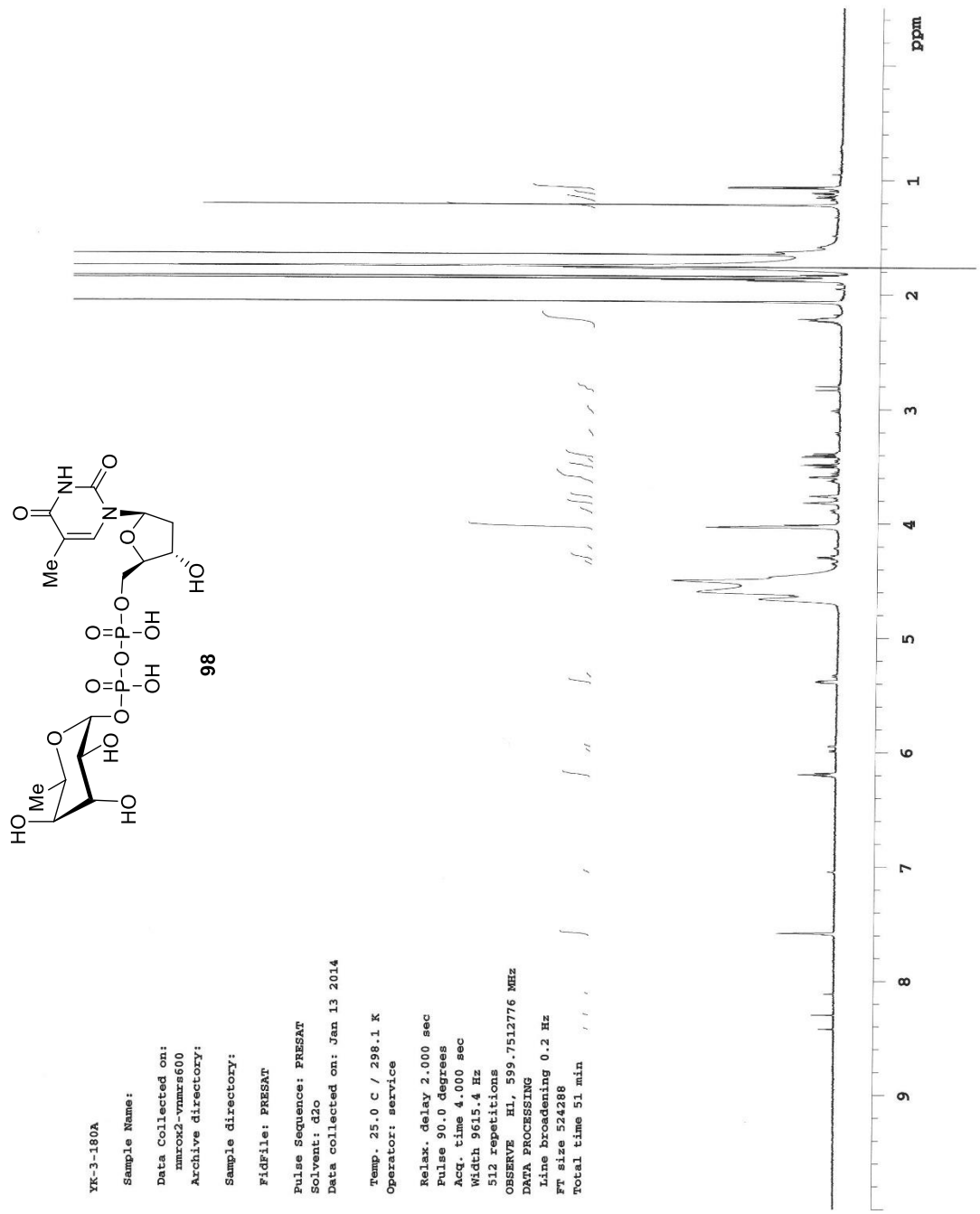


Figure 2-4. ^{1}H NMR of TDP-6-deoxy-D-gulose (**98**).

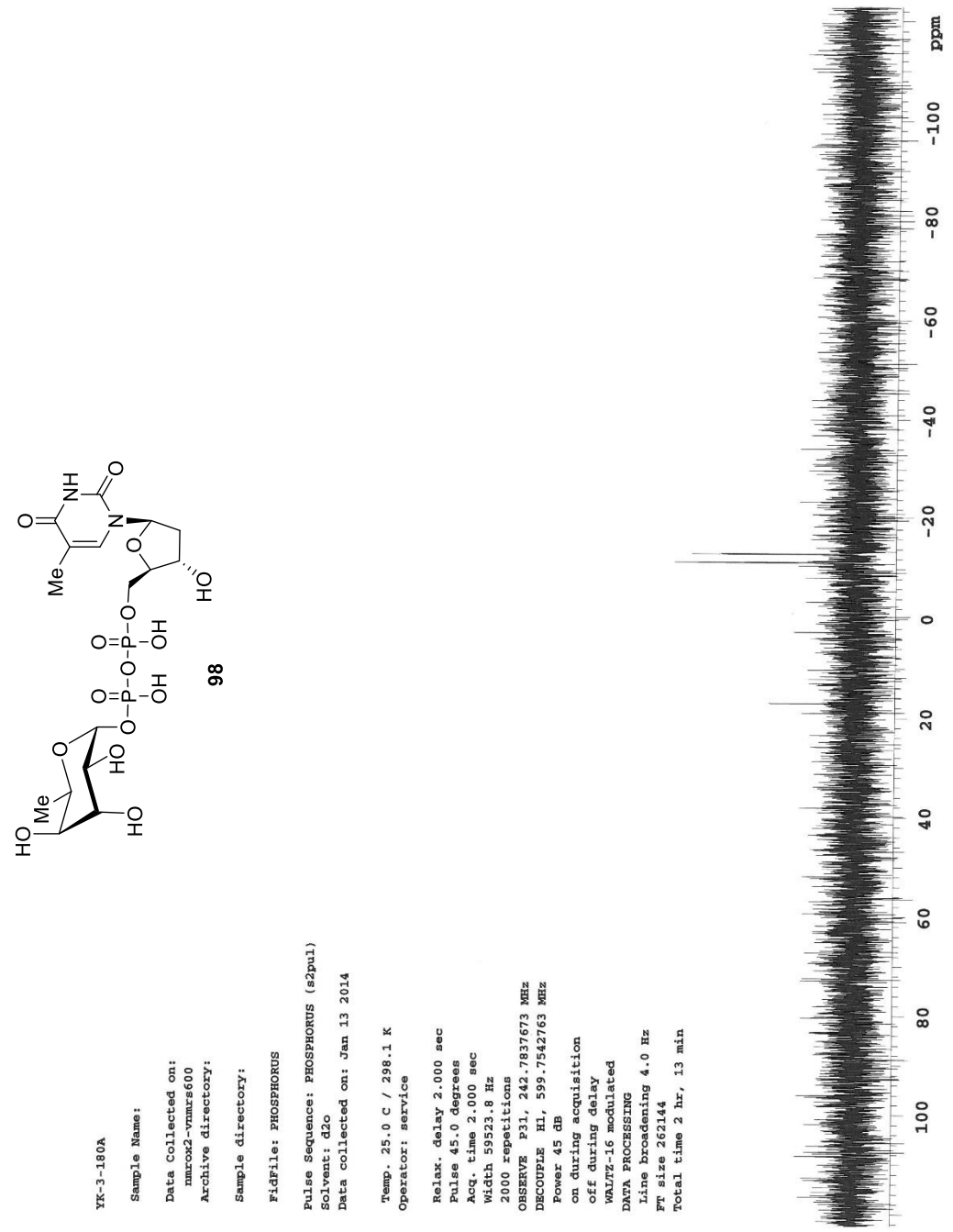


Figure 2-5. ^{31}P NMR of TDP-6-deoxy-D-gulose (**98**).

2.2.3.4 TDP-3-amino-3-deoxy-D-fucose

A 5 mL solution containing 30 mM TDP-D-glucose (**38**), 0.3 mM NAD, 50 mM Tris (pH 7.5), and approximately 0.5 mg RfbB was incubated at 37 °C for 5 h. The reaction mixture was made 10 mM in L-glutamate, 0.8 mM in PLP and approximately 0.3 mg FdtA and 1.5 mg DesV were added. After incubation at 37 °C for 4 h, the reaction mixture was filtered using a Millipore YM-10 ultrafiltration membrane and lyophilized. The product TDP-3-amino-3-deoxy-D-fucose (**101**) was purified by semiprep HPLC in the same manner as **98**, and **101** typically elutes at approximately 6 min. The collected compound was lyophilized to dryness and characterized by ESI-MS and NMR spectroscopy. The observed ^{1}H NMR spectra are consistent with the reported values.[90] HRMS (ESI, negative ion mode) calculated for $C_{16}H_{26}N_3O_{14}P_2^-$ [M – H]$^-$: 546.0895 m/z, found: 546.0882 m/z.

2.2.3.5 Dihydropyran standards

The (2*R*,3*R*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran standard (**105**) was prepared according to a previously described procedure (Figure 2-6).[92] A 3 mL solution containing 5 mM TDP-D-glucose (**38**), 0.3 mM NAD, 10 mM magnesium chloride, 50 mM photassium phosphate (pH 7.5), and 0.5 mg RfbB and 0.1 mg Tyl1a was incubated at 37 °C for 5 h. Then the reaction mixture was filtrated using a Millipore YM-10 ultrafiltration membrane and the filtrate was lyophilized. The product was purified by HPLC in the same manner as **101**, and the desired product typically elutes at approximately 2 min, which is consistent with previous report. The collected compound was lyophilized to dryness and characterized by ESI-MS. HRMS (ESI, positive ion mode) calculated for $C_6H_9O_4^+$ [M + H]$^+$: 145.0495 m/z, observed: 145.0494 m/z. The (2*R*,3*R*)-2-methyl-3-hydroxy-4-keto-2,3-dihydropyran standard (**106**) was chemically synthesized by Geng-Min Lin (*manuscript in preparation*).

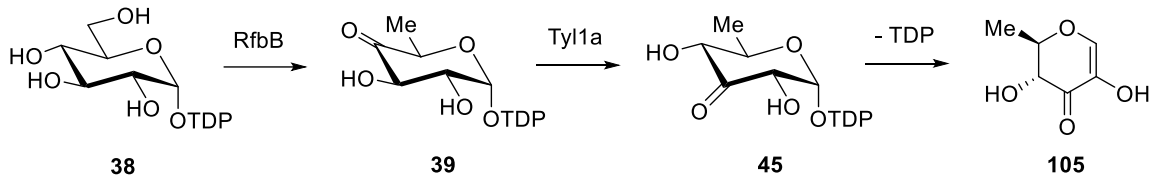


Figure 2-6. Preparation of (2*R*,3*R*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran (**105**).

2.2.4 DesII and DesII/DesV Assays

A typical reaction of DesII was conducted by incubating 0.25–1 mM substrate and 0.5–1 mM SAM with 10–20 μM DesII in an anaerobic buffer containing 25 mM EPPS (pH 8.0), 1 mM DTT and 0.5–1 mM $Na_2S_2O_4$ at room temperature or at 30 °C. When the incubation was longer than 2 h, additional $Na_2S_2O_4$ was added every 2 h to compensate for its decomposition. Reaction aliquots of 50 or 100 μL were taken at appropriate times and deproteinized using YM-10 centrifugal filters. The filtrate of each sample was then frozen and stored at –80 °C until analyzed by HPLC. An analytical 4 × 25 mm Dionex CarboPac PA1 anion exchange HPLC column was employed using a two-solvent system: solvent A was H_2O, and solvent B was 1.5 M ammonium acetate in H_2O (pH 7.0). The flow rate was 1 mL/min and the UV-absorbance detector was set to 267 nm. The column was eluted with a two-phase linear gradient: 2.5 to 15% B in 20 min, 15 to 20% B in 20 min and then return to 2.5% B in 5 min. For the coupled DesII and DesV assay, 10 mM L-glutamate, 0.8 mM PLP and 10 μM DesV were included in the reaction mixture. Work up and analysis of reactions that included the DesV assay components were the same as those described above for DesII alone.

Reduction with $NaBH_4$ was used to trap the TDP-sugar product(s) containing the 3-keto functional group. $NaBH_4$ was added to the reaction aliquots taken at given time

points to a final concentration of 100 mM. After additional incubation on ice for 1 h, the reaction mixture was analyzed by HPLC as described above. In some cases, 5 mM $NaBH_4$ was included in the reaction mixture at the beginning of the reaction, and the reaction mixture was analyzed by HPLC at given time points. The enzyme activity in the presence or absence of 5 mM $NaBH_4$ was found to be the same.

The formation of dihydropyrans such as **105** and **106** in the DesII reaction with TDP-[3-^{2}H]-6-deoxy-D-mannose (**87D**) was analyzed by employing an analytical 250 × 4.6 mm Agilent Microsorb 100-5 C18 HPLC column in a two-solvent system of A, 1% ammonium acetate (pH 5.1), and B, acetonitrile. The samples were eluted with 2% B for 5 min, 2 to 10% B for 18 min, 10 to 35% B for 9 min, and then return to 2% B for 2 min. The flow rate was 1 mL/min and the detector was set to 267 nm.

2.2.5 ^{1}H NMR Characterization of Species Y

Three preparatory 5 mL reaction mixtures containing 0.25–0.5 mM SAM, 0.12–0.250 mM **86**, 0.5 mM $Na_2S_2O_4$, and 10–15 μM DesII in 25 mM EPPS (pH 8.0) were incubated anaerobically at room temperature for 8 h, and additional $Na_2S_2O_4$ was added every 2 h. The reaction mixture was deproteinized using a Millipore YM-10 ultrafiltration membrane and lyophilized. The product **Y** was isolated using a Dionex CarboPac PA1 semiprep column with a flow rate of 3 mL/min. The column was eluted using a two solvent system of A, H_2O, and B, 1 M ammonium bicarbonate in H_2O, with the following gradient: 3.75 to 30% B in 40 min, followed by a return to 3.75% B in 5 min. The collected product was lyophilized to dryness prior to analysis by ESI-MS and 1-D as well as 2-D ^{1}H NMR spectroscopy. Two contaminating species were also observed in the NMR spectra in addition to TDP-6-deoxy-D-gulose: a 1,2-cyclic sugar phosphate and thymidine

monophosphate (TMP). ^{1}H NMR and ^{31}P NMR spectra are shown in Figure 2-7 and Figure 2-8, respectively.

Species **Y**: ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 7.60 (s, 1H, 6-H), 6.19–6.24 (m, 1H, 1′-H, convoluted with 1′-H of TMP), 5.41 (dd, 1H, J = 3.7, 7.2 Hz, 1″-H), 4.46–4.51 (m, convoluted with HOD peak, 1H, 3′-H), 4.29–4.34 (m, 1H, 5″-H), 4.05 (d, 3H, J = 4.4 Hz, 4′-H, 5′-H), 3.84 (dd, 1H, J = 3.7, 3.7 Hz, 3″-H), 3.78 (ddd, 1H, J = 3.7, 3.7, 3.7 Hz, 2″-H), 3.62 (d, 1H, J = 3.7 Hz, 4″-H), 2.18–2.31 (m, 2H, 2′-H, convoluted with 2′-H of TMP), 1.80 (s, 3H, CH_3, convoluted with CH_3 of TMP), 1.09 (d, 3H, J = 6.0 Hz, 6″-H); ^{31}P NMR (D_2O, 243 MHz) δ (ppm) –11.3 (d, J = 23 Hz), –13.0 (d, J = 20 Hz). HRMS (ESI, negative ion mode) calculated for $C_{16}H_{25}O_{15}N_2P_2^-$ $[M - H]^-$: 547.0736 m/z, found: 547.0729 m/z.

Contaminant 1,2-cyclic sugar-phosphate: ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 5.59 (dd, 1H, J = 5.2, 15.8 Hz, 1-H), 4.60 (dd, 1H, J = 2.9, 5.2 Hz, 2-H), 4.29–4.34 (m, 1H, 5-H), 3.69 (dd, 1H, J = 4.7, 6.6 Hz, 4-H), 3.64–3.68 (m, 1H, 3-H), 1.10 (d, 3H, J = 6.0 Hz, 6-H); ^{31}P NMR (D_2O, 243 MHz) δ (ppm) 16.18 (s). HRMS (ESI, negative ion mode) calculated for $C_6H_{10}O_7P^-$ $[M - H]^-$: 225.0170 m/z, found: 225.0175 m/z.

Contaminant TMP: ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 7.67 (s, 1H, 6-H), 6.19–6.24 (m, 1H, 1′-H, convoluted with 1′-H of **Y**), 4.42–4.47 (m, convoluted with HOD peak, 1H, 3′-H), 3.98–4.02 (m, 1H, J = 3.7 Hz, 4′-H), 3.81 (dd, 2H, J = 3.7, 3.7 Hz, 5′-H), 2.18–2.31 (m, 2H, 2′-H, convoluted with 2′-H of **Y**), 1.80 (s, 3H, CH_3, convoluted with CH_3 of **Y**); ^{31}P NMR (D_2O, 243 MHz) δ (ppm) 3.63 (s). HRMS (ESI, negative ion mode) calculated for $C_{10}H_{14}O_8N_2P^-$ $[M - H]^-$: 321.0493 m/z, found: 321.0496 m/z.

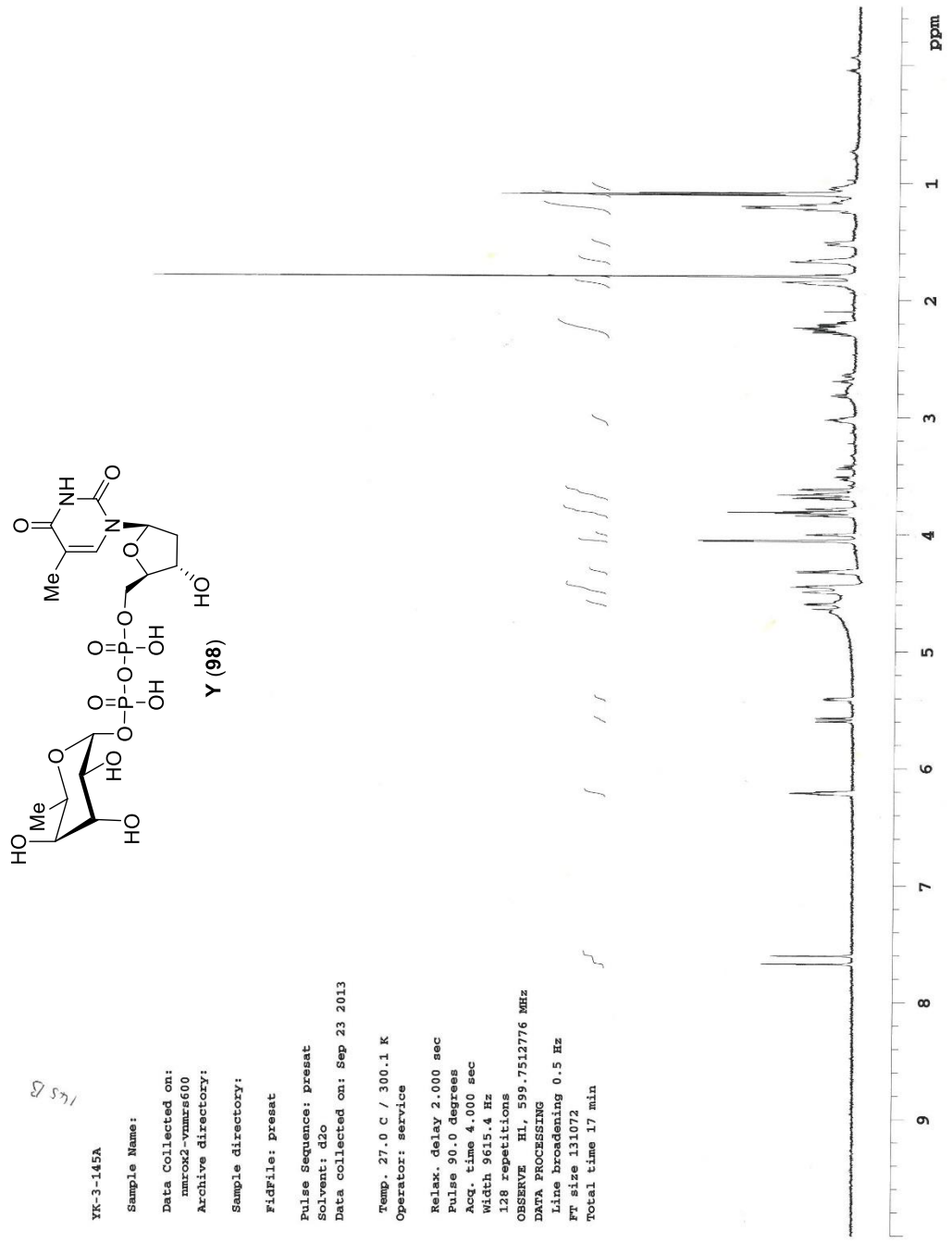


Figure 2-7. ^{1}H NMR of **Y**.

Y (98)

Figure 2-8. ^{31}P NMR of **Y**.

2.2.6 Deuterium Incorporation Experiment with TDP-D-fucose

Anaerobic D_2O was prepared via three freeze-thaw cycles using an anaerobic Schlenk line supplied with an argon atmosphere. Buffer containing 25 mM EPPS (pH 8.0) was prepared in H_2O, dried under vacuum, reconstituted with D_2O, dried again under vacuum and transferred to the anaerobic glovebox where it was reconstituted with the anaerobic D_2O. Freshly reconstituted DesII was exchanged with the anaerobic EPPS buffer in D_2O via repeated cycles of concentration using YM-10 centrifugal filters followed by addition of D_2O buffer. Reactions were run by incubating approximately 10 μM DesII in the D_2O EPPS buffer with 0.5 mM SAM, 0.25 mM TDP-D-fucose (**86**) and 0.5 mM $Na_2S_2O_4$ for 2.5–4 h. The reaction mixture was deproteinized using YM-10 centrifugal filters and the residual TDP-D-fucose (**86**) and newly formed TDP-6-deoxy-D-gulose (**98**) in the filtrate were isolated by analytical HPLC as described above. The collected samples were lyophilized to dryness and analyzed by ESI-MS.

The C3-radical intermediate of TDP-D-fucose (*i.e.*, **97** with concentration [**97**]) is treated as decomposing in different ways, each governed by a pseudo-first order rate constant under the experimental conditions (see Figure 2-19). Intermediate **97** can partition to regenerate TDP-D-fucose (**86** with concentration [**86**]) in two ways: either the reverse of **97** formation, or net H-atom transfer to the *re*-face of **97** (governed by k_{re}). In D_2O containing buffers, TDP-D-[3-^{2}H]fucose (**86D** with concentration [**86D**]) will also be formed, but only via the latter pathway with net D-atom transfer to **97**. In contrast TDP-6-deoxy-D-gulose (**98** with concentration [**98**]) is only formed via net H-atom transfer to the *si*-face of **97** (governed by k_{si}). Likewise, TDP-6-deoxy-D-[3-^{2}H]gulose (**98D** with concentration [**98D**]) will also be formed in D_2O containing buffers. If we consider a reaction run in a D_2O containing buffer with mole fractions f_H and f_D of solvent exchangeable H- and D-atoms, then we can write the following rate equations

$$d\,[\mathbf{86D}]/dt = k_{re}\,f_D[\mathbf{97}] \,/\, {}^{D}k_{re}, \qquad [\mathbf{1a}]$$

$$d\,[\mathbf{98D}]/dt = k_{si}\,f_D[\mathbf{97}] \,/\, {}^{D}k_{si}, \qquad [\mathbf{1b}]$$

where ${}^{D}k_{re}$ and ${}^{D}k_{si}$ are the solvent deuterium kinetic isotope effects (KIEs) on k_{re} and k_{si}, respectively. It should be pointed out that reaction of **86D** with free enzyme to form **97** is considered insignificant in this treatment. This will be a reasonable omission, if the primary deuterium KIE on H-atom abstraction from C3 is large and [**86**] is much greater than [**86D**], which was borne out by the experiments. In this system of differential equations, [**86D**] and [**98D**] change simultaneously so that Eqns. [**1a**] and [**1b**] can be combined to eliminate *dt* providing

$$d\,[\mathbf{86D}]/d\,[\mathbf{98D}] = k_{re}\,{}^{D}k_{si} \,/\, (k_{si}\,{}^{D}k_{re}). \quad [\mathbf{2}]$$

If the solvent deuterium KIEs on net H-atom transfer to the *re*- and *si*-faces of **97** are similar (*i.e.*, ${}^{D}k_{re} \approx {}^{D}k_{si}$), then Eqn. [**2**] can be simplified to

$$d\,[\mathbf{86D}]/d\,[\mathbf{98D}] = k_{re} \,/\, k_{si}. \qquad [\mathbf{3}]$$

Because the initial concentrations of **86D** and **98D** are equal to 0, expression [**3**] implies that the ratio [**86D**]/[**98D**] is a constant versus time and expresses the relative partitioning of **97** between net H-atom transfer to the *re*- versus *si*-faces of the C3-radical (*i.e.*, k_{re} / k_{si}).

We can use HPLC peak integrations to measure the fractional concentrations of all isotopologs of TDP-D-fucose (denoted $f_{\mathbf{86}}$) and all isotopologs TDP-6-deoxy-D-gulose (denoted $f_{\mathbf{98}}$) at some reaction time point. Likewise, we can use mass spectrometry to measure the corresponding fractional deuteration of TDP-D-fucose (denoted $\beta_{\mathbf{86}}$) and TDP-6-deoxy-D-gulose (denoted $\beta_{\mathbf{98}}$). These observables then permit calculation of the ratio [**86D**]/[**98D**] for a given sample according to

$$[\mathbf{86D}]/\,[\mathbf{98D}] = f_{\mathbf{86}}\,\beta_{\mathbf{86}} \,/\, (f_{\mathbf{98}}\,\beta_{\mathbf{98}}), \qquad [\mathbf{4}]$$

which determines k_{re} / k_{si} as discussed above. Finally, we can use this ratio to determine the relative preference χ_{re} of H-atom transfer to the *re*-face, which we define quantitatively as

$$\chi_{re} := (k_{re} / k_{si}) / (1 + k_{re} / k_{si}). \qquad [\mathbf{5}]$$

A value of 0 or 1 for χ_{re} would indicate stereospecific H-atom transfer to the *si*- or *re*-face, respectively, whereas a value of 0.5 would indicate no stereochemical preference of one face over the other.

In order to obtain the fractional deuterations $f_{\mathbf{86}}$ and $\beta_{\mathbf{86}}$, the peaks corresponding to TDP-D-fucose and TDP-6-deoxy-D-gulose were isolated by HPLC and analyzed by ESI-MS. The ratio R_M of the mass spectroscopic ion signal intensities for a given TDP-sugar, *i.e.*,

$$R_M := (\text{intensity of } [M + 1]) / (\text{intensity of } [M]), \qquad [\mathbf{6}]$$

must be corrected for natural abundance heavy isotopes. It is straightforward to show that this can be accomplished by subtracting from R_M a term α which is just the ratio R_M when there is no labeling. Applying the correction term, we then have an expression for the determination of β for either TDP-sugar,

$$\beta = (R_M - \alpha) / (R_M - \alpha + 1). \qquad [\mathbf{7}]$$

The correction term α can also be calculated based on the 1.11% natural abundance of ^{13}C providing $\alpha = 0.2$,[1] which will be the same for the two isomers TDP-D-fucose and TDP-6-deoxy-D-gulose. This calculated value was not significantly different from the measured value of 0.18 ± 0.02, which also offers some insight into the precision by which the R_M ratios could be measured in these experiments.

Three experimental trials were attempted to measure χ_{re} as shown in Table 2-1. While the result from the first trial suggested preferential net H-atom transfer to the *si*-face, this was not repeatable in the subsequent trials. It should be noted that this experiment was complicated by a solvent deuterium KIE that favored dehydration and/or dehydrogenation

[1] Contributions from natural abundance ^{2}H, ^{15}N and ^{17}O were not significant in light of the measurement precision.

over H-atom transfer in D_2O containing buffers. Thus, even at the nearly 80% observable reaction in Trial III, only approximately 10% deuteration of the residual TDP-D-fucose was expected assuming no stereoselectivity. The mean value of χ_{re} from the three trials was 0.3 ± 0.1 (± standard error, standard deviation is 0.2). Therefore, net H-atom transfer to the C3 radical does not appear to be stereospecific, though there may be a small selectivity for the *si*-face.

Table 2-1. Results from three trials to measure relative partitioning of H-atom transfer to the *re*- versus *si*-face of the C3 radical of TDP-D-fucose during reaction with DesII. The correction factor α for natural abundance ^{13}C was 0.2 in all cases. $R_{M,\mathbf{86}}$ and $R_{M,\mathbf{98}}$ are the respective R_M ratios measured for TDP-D-fucose and TDP-6-deoxy-D-gulose post reaction. Values in parentheses for $R_{M,\mathbf{86}} - \alpha$ represent the expected values for this measurement in the given experiment assuming $\chi_{re} = 0.5$ (*i.e.*, no stereoselectivity of net H-atom transfer). All values are listed arbitrarily to three decimal places.

Parameter	Trial I	Trial II	Trial III
$f_{\mathbf{86}}$	0.432	0.532	0.227
$f_{\mathbf{98}}$	0.080	0.080	0.029
$R_{M,\mathbf{86}} - \alpha$	0.009 (0.130)	0.039 (0.054)	0.095 (0.115)
$R_{M,\mathbf{98}} - \alpha$	1.669	0.507	4.222
$f_{\mathbf{86}}\beta_{\mathbf{86}} / f_{\mathbf{98}}\beta_{\mathbf{98}}$	0.082	0.738	0.841
χ_{re}	0.075	0.425	0.457

2.2.7 Radical Trap Experiments with TDP-D-fucose

The DesII reaction was also performed in the presence of the radical trap 5,5-dimethyl-1-pyrroline-*N*-oxide (DMPO) to further interrogate the solvent accessibility of the substrate radical derived from TDP-D-fucose. 15 μM DesII was incubated at room temperature for 6 h in a reaction mixture containing 0.5 mM SAM, 0.25 mM TDP-D-fucose (**86**), and 0.5 mM $Na_2S_2O_4$ in the presence and absence of 100 mM DMPO. Each sample was analyzed by HPLC, and any new peaks were collected, lyophilized and analyzed by ESI-MS.

2.3 Results and Discussion

2.3.1 DesII Reaction with TDP-4-amino-4-deoxy-D-fucose (85)

TDP-4-amino-4-deoxy-D-fucose (**85**) is a C4 epimer of the natural substrate of DesII (*i.e.,* TDP-4-amino-4,6-dideoxy-D-glucose (**40**)), which was designed to test the effect of the stereochemical change at C4. Compound **85** was chemically synthesized with a deuterium introduced at C3 (**85D**, see Figure 2-11) in order to test for H-atom abstraction at C3 when the stereochemical configuration at C4 is altered. The consumption of both **85D** and SAM was observed upon incubation with DesII (Figure 2-9). The formation of 5′-deoxyadenosine (5′-dAdo) was indicative of the reaction progress, and the incorporation of deuterium into 5′-dAdo was observed by MS consistent with C3-centered radical generation (Figure 2-10, A; calculated for $C_{10}H_{11}DO_3N_5^-$ $[M - H]^-$: 251.1008 *m*/*z*, found: 251.1 *m*/*z*). In addition to 5′-dAdo, a new product was observed by HPLC at a retention time of 26.9 min. The new product co-eluted with the deamination product **41** from the

DesII reaction with **40**, indicating that DesII can recognize **85** as a substrate and can catalyze the deamination reaction (see Figure 2-11). The specific activity of DesII toward **85** was not measured. No other TDP-sugar product was observed by HPLC. The level of deuteration of residual **85D** determined by MS remained the same after incubation with DesII for 6 h (ca. 95%, Figure 2-10, B; calculated for $C_{16}H_{25}DO_{14}N_3P_2^-$ [M – H]$^-$: 547.0958 m/z, found: 547.1 m/z; calculated for $C_{16}H_{24}DO_{14}N_3P_2Na^-$ [M – 2H + Na]$^-$: 569.0778 m/z, found: 569.1 m/z), suggesting that there was no hydrogen/deuterium exchange at C3 of **85**.

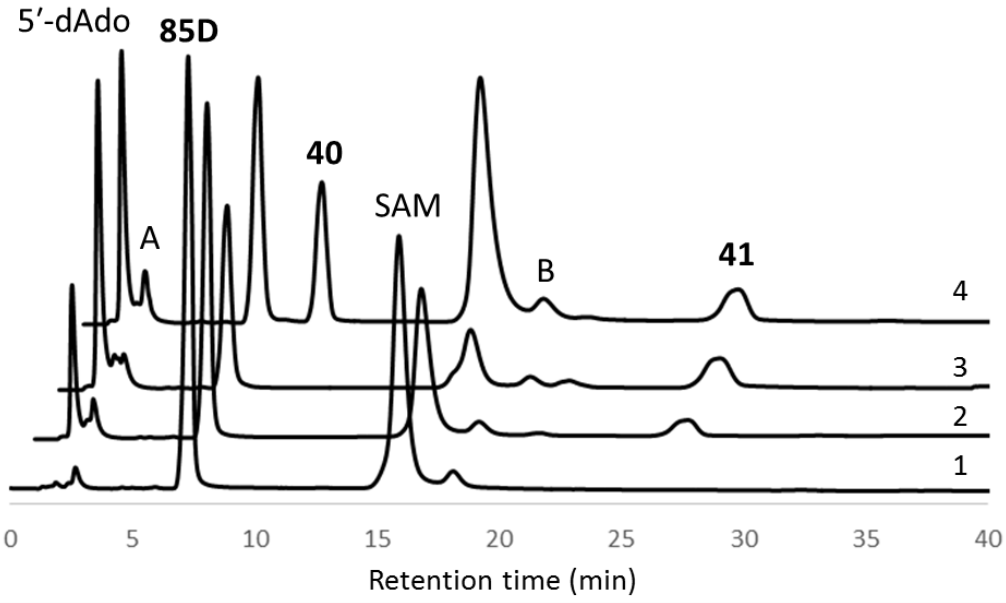


Figure 2-9. HPLC traces showing the consumption of TDP-[3-^{2}H]-4-amino-4-deoxy-D-fucose (**85D**) and SAM in the presence of DesII and $Na_2S_2O_4$. Trace 1 was measured after 30 min without DesII. Traces 2 and 3 were measured after 2 h and 6 h in the presence of DesII, respectively. Trace 4 was the same as trace 2 coinjected with a DesII reaction with TDP-4-amino-4,6-dideoxy-D-glucose (**40**) for 30 min. The retention times in traces 2, 3, and 4 are shifted by 1, 2, and 3 min compared to trace 1. Peak A corresponds to the methylthioadenosine (MTA) decomposition product of SAM.[94] Peak B corresponds to a contaminant in the SAM reagent.

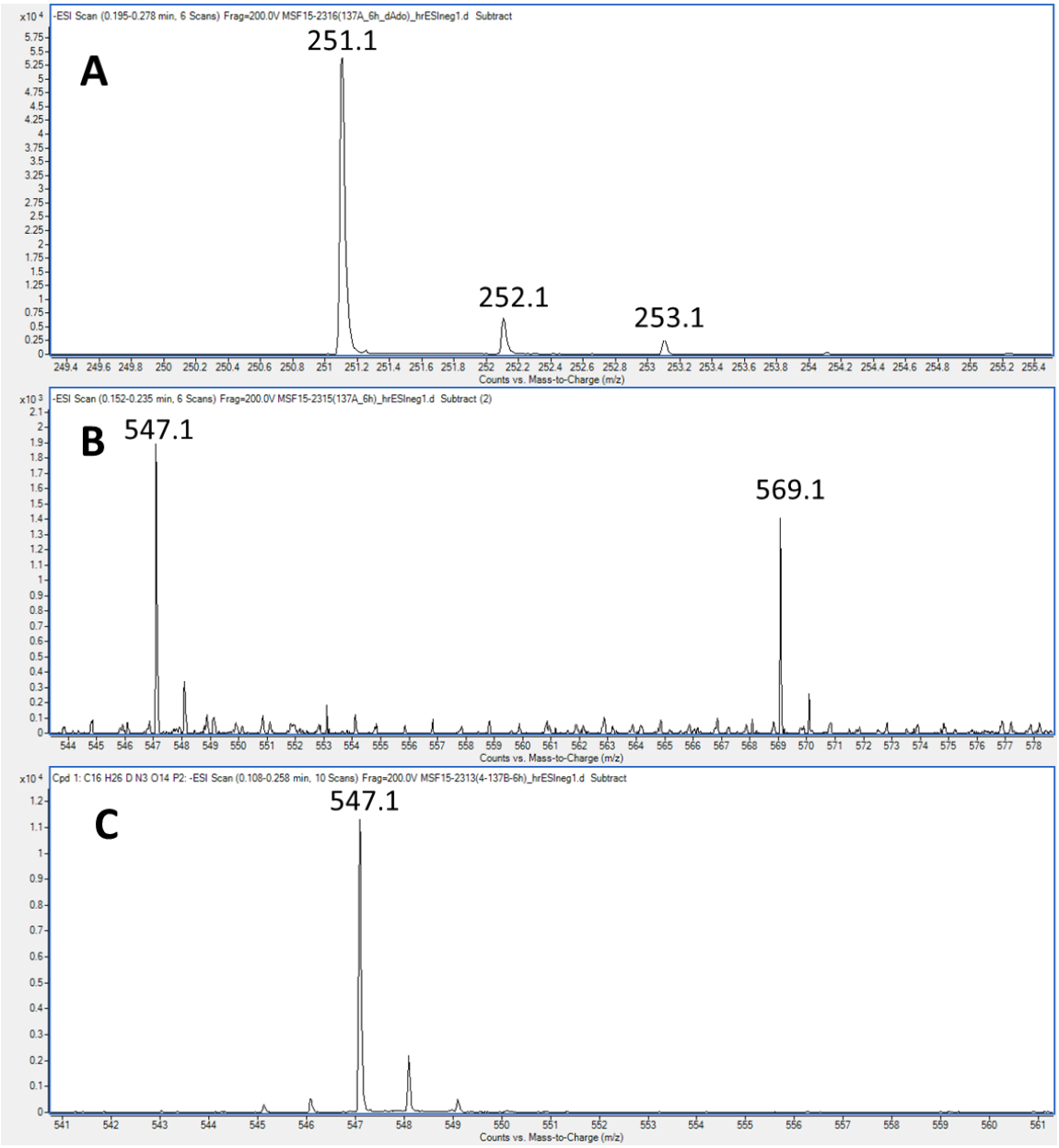


Figure 2-10. ESI-MS (negative ion mode) of A: 5′-dAdo generated from the DesII reaction with **85D**, B: residual **85D**, C: residual **85D** after reduction by $NaBH_4$.

Although it seemed that DesII does not catalyze the dehydrogenation of **85** as shown by incubating **85** with DesII alone, this activity could not be completely ruled out, because the dehydrogenation product (**88**, see Figure 2-11), if formed, is expected to be unstable. Thus, the possibility of formation of **88** was further investigated by including $NaBH_4$ in the reaction mixture. If the dehydrogenation reaction takes place in the presence of DesII to generate **88**, the addition of $NaBH_4$ would result in the regeneration of undeuterated **85**. The remaining substrate **85D** from the DesII/$NaBH_4$ reaction mixture incubated for 6 h was isolated by HPLC for MS analysis, which was shown to be 95% deuterated (Figure 2-10, C; calculated for $C_{16}H_{25}DO_{14}N_3P_2^-$ [M – H]$^-$: 547.0958 m/z, found: 547.1 m/z). No decrease in the level of deuteration was consistent with no or negligible formation of **88**.

All of the results indicated that TDP-4-amino-4-deoxy-D-fucose (**85**) with the axial amino substituent at C4 is processed by DesII as a substrate (Figure 2-11). **85** undergoes a deamination reaction as does the natural substrate **40**. The substrate radical generated by abstraction of hydrogen at C3 proceeds only to the deamination pathway, and no other reaction product such as dehydrogenation or reduction of substrate radical was observed.

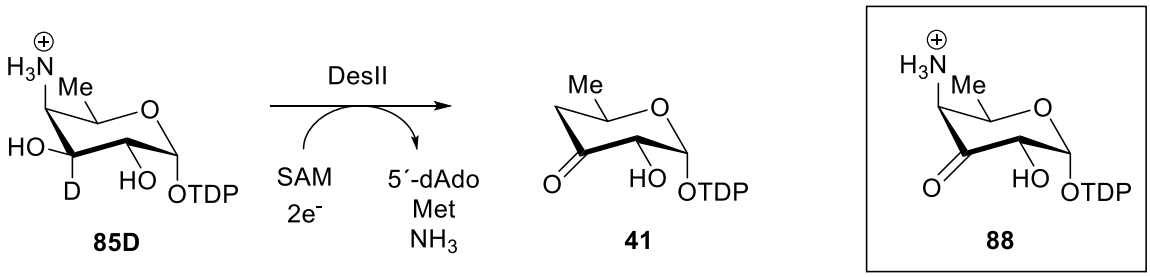


Figure 2-11. DesII-catalyzed deamination of **85D**. The dehydrogenation product **88** was not observed.

2.3.2 DesII Reaction with TDP-D-fucose (86)

TDP-D-fucose (**86**), the C4 epimer of TDP-D-quinovose (**44**), was employed as a potential substrate of DesII to test the hypothesis that a stereochemical change that may permit better hyperconjugation could result in dehydratase activity. Previously, it was reported that TDP-D-fucose (**86**) was not a substrate for DesII.[40] However, in light of improvements in both reconstitution and handling of the enzyme since that initial report, we decided to revisit the question.

TDP-D-fucose (**86**) was prepared chemoenzymatically via C6-dehydration of TDP-D-glucose (**38**) using the 4,6-dehydratase RfbB followed by reduction of the TDP-6-deoxy-4-keto-D-glucose product (**39**) with $NaBH_4$.[40] The consumption of both **86** and SAM and the production of 5′-dAdo were only observed upon prolonged incubation (hour time scale) with a high concentration of DesII (ca. 10–20 μM) in the presence of $Na_2S_2O_4$. This corresponded to a specific activity on the order of 5×10^{-4} $\mu mol \cdot min^{-1} mg^{-1}$ for consumption of **86** versus 1 $\mu mol \cdot min^{-1} mg^{-1}$ for dehydrogenation of **44** at pH 8.0.[43] The much lower specific activity of DesII towards **86** helps to explain why it was not previously recognized as a substrate.[40]

The HPLC analysis of the reaction of DesII with **86** showed that three new products, designated as **X**, **Y**, and **Z** in the order of retention time, were generated in addition to 5′-dAdo (Figure 2-12). The species **X** had the earliest retention time of 28.4 min. It was found that species **X** was also generated during prolonged incubation of DesII with SAM and $Na_2S_2O_4$ in the absence of **86**, while there was no significant formation of 5′-dAdo (Figure 2-13). If either SAM, DesII or $Na_2S_2O_4$ was excluded from the reaction mixture, however, species **X** was not formed, suggesting that species **X** may be a sulfinate adduct of deoxyadenosine. The ESI-MS analysis of species **X** showed signals at *m*/*z* 316.1 (positive ion **X**, Figure 2-14, A) and *m*/*z* 314.1 (negative ion **X**, Figure 2-14, B). This is consistent

with the formation of 5′-deoxyadenosine-5′-sulfinic acid, which has a calculated monoisotopic mass of 315.1 Da. When the dideuterated isotopolog [5′,5′-$^{2}H_2$]SAM was used instead of natural abundance SAM, ESI-MS peaks at *m*/*z* 318.1 (positive ion $\mathbf{X_d}$, Figure 2-14, C) and 316.1 *m*/*z* (negative ion $\mathbf{X_d}$, Figure 2-14, D) were observed. Moreover, a species with the same HPLC retention time as **X** could be produced from a photoirradiated solution of adenosylcobalamin and $Na_2S_2O_4$ (unpublished data). Based on these observations, species **X** was assigned as 5′-deoxyadenosyl-5′-sulfinic acid (**89**) generated from SAM (Figure 2-15). The specific activity for the sulfination reaction with 500 μM $Na_2S_2O_4$ appeared to be no greater than 2×10^{-3} μmol·min^{-1}mg^{-1} in the absence of **86**, and its rate of formation was reduced in the presence of the TDP-sugar. This suggests that dithionite can access the DesII active site in the absence of a TDP-sugar and intercept the 5′-deoxyadenosyl radical either following or concerted with reductive homolysis of SAM.

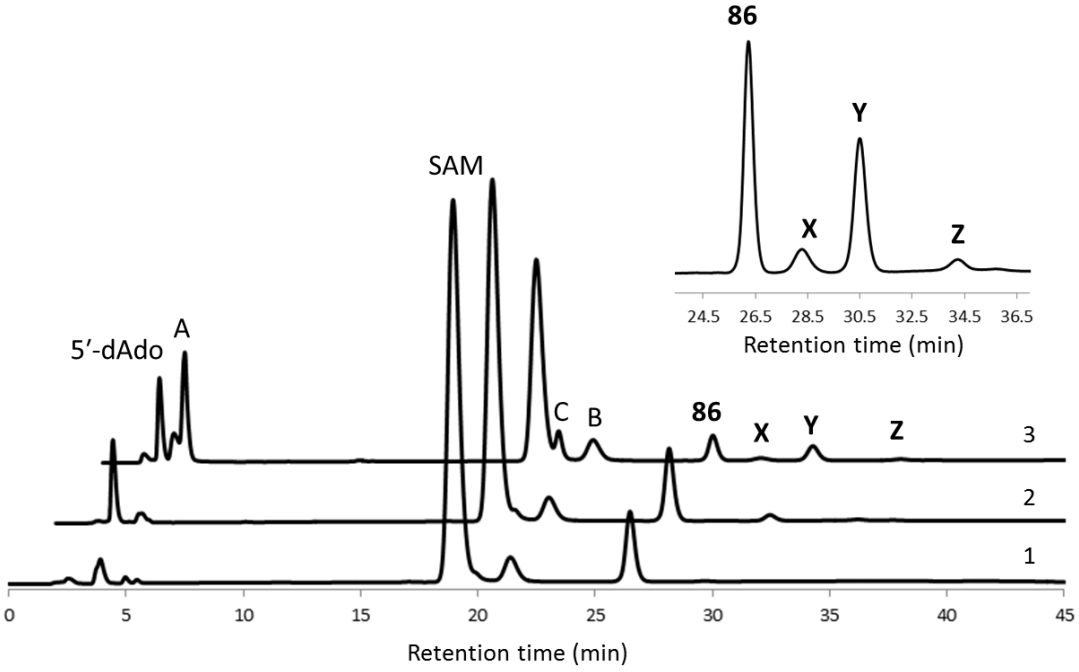


Figure 2-12. HPLC traces showing the reaction of DesII with TDP-D-fucose (**86**). Trace 1 was measured after 2 h without DesII. Traces 2 and 3 were measured after 30 min and 6 h in the presence of DesII. The retention times in traces 2 and 3 are shifted by 2 and 4 min compared to trace 1. Three new products (**X**, **Y**, and **Z**) are observed as shown in the inset (from trace 3). Peaks A and B correspond to MTA and the contaminant in the SAM reagent, respectively. Peak C corresponds to thymidine monophosphate (TMP).

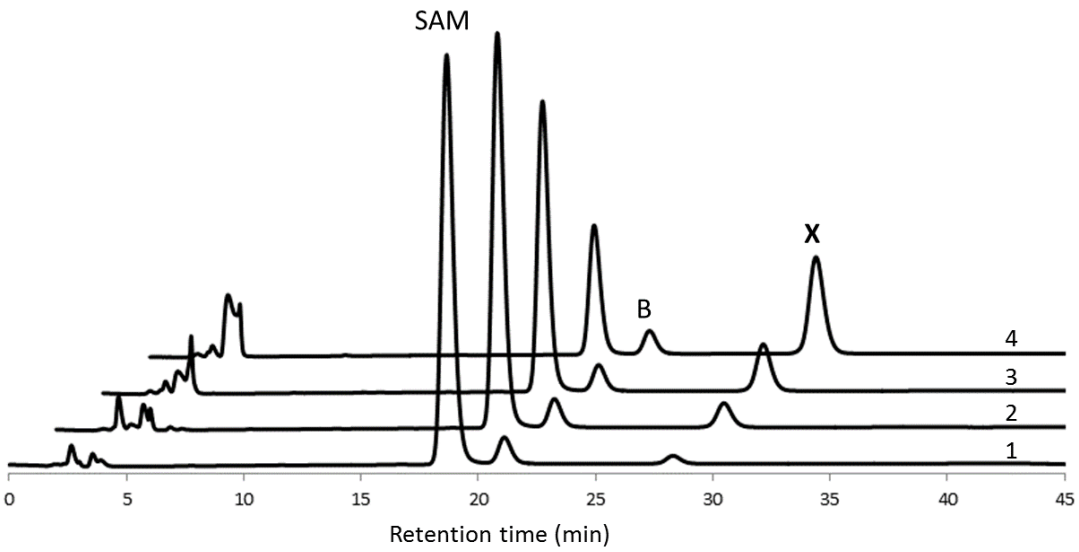


Figure 2-13. HPLC traces showing the consumption of SAM by DesII in the presence of $Na_2S_2O_4$. Traces 1, 2, 3, and 4 were measured after 30 min, 2 h, 6 h, and 24 h, respectively. The retention times in traces 2, 3, and 4 are shifted by 2, 4, and 6 min compared to trace 1. The formation of **X** is noted.

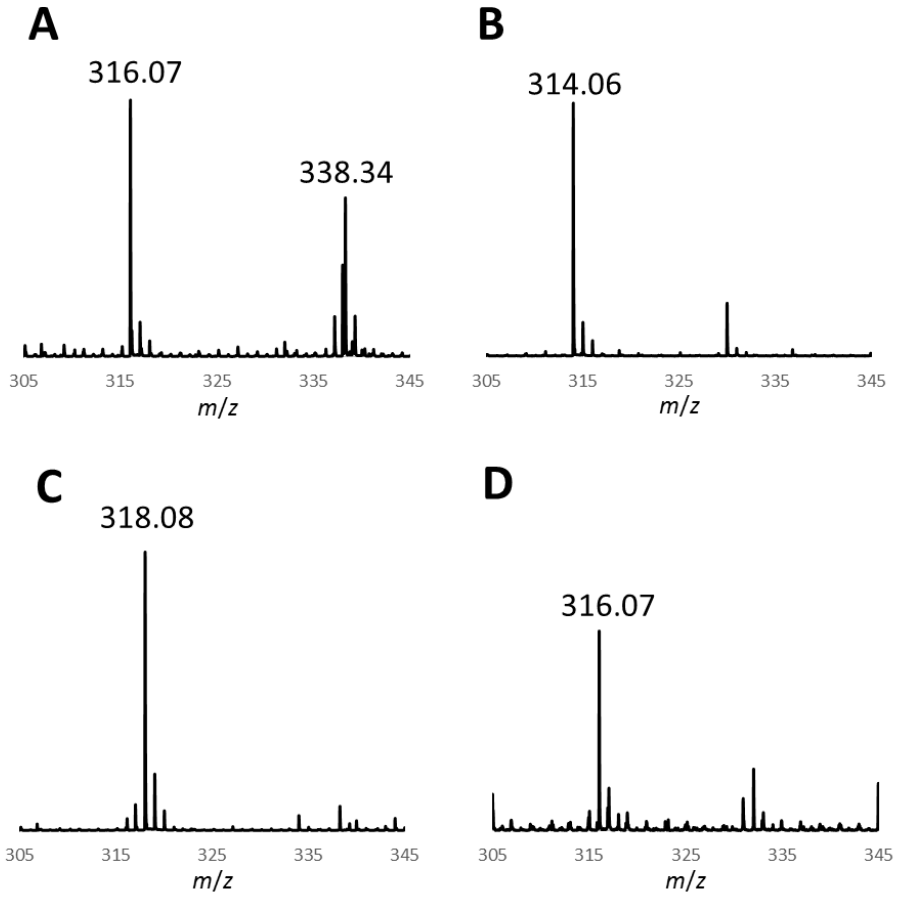


Figure 2-14. ESI-MS of A: **X** (positive ion), B: **X** (negative ion), C: $\mathbf{X_d}$ (positive ion), and D: $\mathbf{X_d}$ (negative ion). The peak at *m/z* 338.34 likely corresponds to a known contaminant (*i.e.* erucamide) in the ESI-MS.

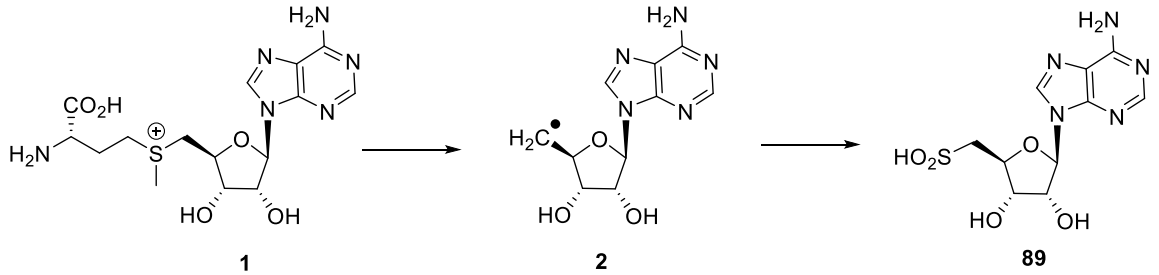


Figure 2-15. Conversion of SAM (**1**) to 5′-deoxyadenosyl-5′-sulfinic acid (**89**).

The formation of sulfinate adducts has been reported in the reactions of two other radical SAM enzymes when $Na_2S_2O_4$ is used to maintain a reduced $[4Fe\text{-}4S]^{1+}$ cluster. The spore photoproduct lyase (SPL) is a radical SAM enzyme that repairs the thymidine dimer in DNA (*e.g.*, **90** which is a model compound) generated by UV irradiation to two intact thymidines (**93**).[97] A single mutation of a conserved cysteine 141 of SPL resulted in the production of a sulfinate derivative of dithymidine monophosphate (**94**, see Figure 2-16).[98] The atypical radical SAM enzyme Dph2 in diphthamide biosynthesis, which employs a SAM-derived 3-amino-3-carboxypropyl (ACP) radical (**95**) rather than a 5′-deoxyadenosyl radical, has also been shown to produce a sulfinic acid derivative of ACP (**96**, homocysteine sulfinic acid) in the absence of substrate (Figure 2-17).[99] However, the formation of **89** from SAM in the presence of DesII and $Na_2S_2O_4$ represents the first report of trapping of 5′-deoxyadenosyl radical generated by a radical SAM enzyme.

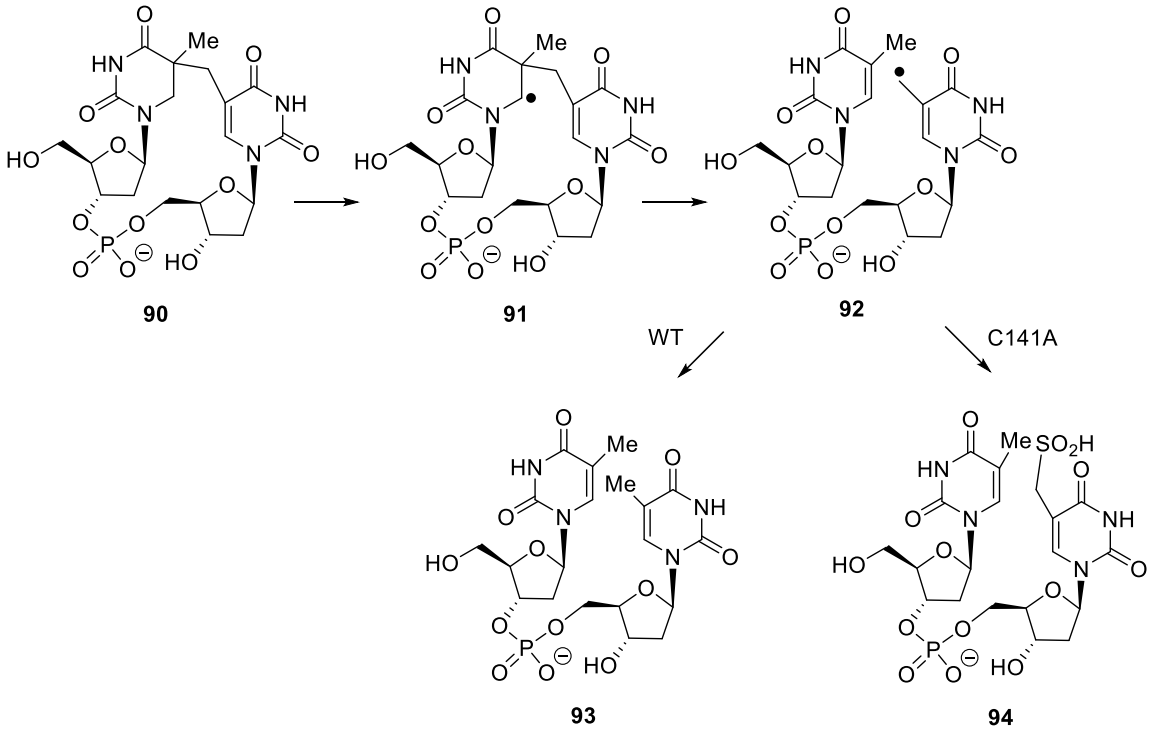


Figure 2-16. Repair of thymidine dimer (**90**) by SPL and the formation of sulfinic acid adduct of dithymidine monophosphate (**94**) by a mutant of SPL.

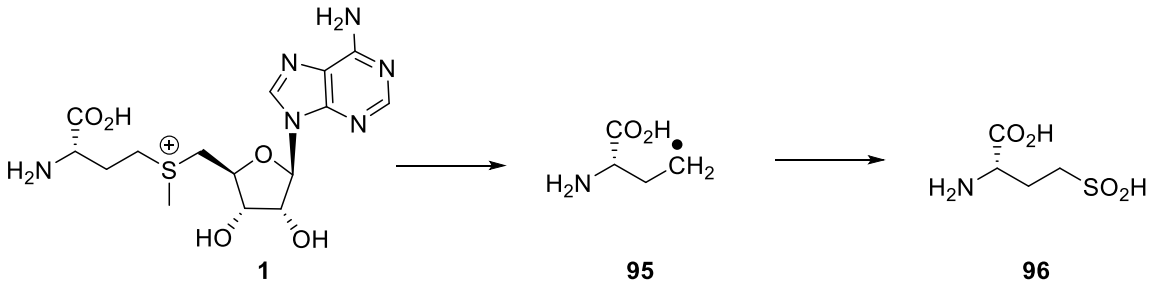


Figure 2-17. Formation of homocysteine sulfinic acid (**96**) in the absence of substrate by Dph2.

Both of the remaining two product peaks in Figure 2-12 originated from TDP-D-fucose (**86**). The major product peak (*i.e.*, **Y** in H_2O buffer) from the DesII reaction with TDP-D-fucose had a retention time of 30.5 min. Negative ion ESI-MS analysis of the isolated **Y** showed clear signals indicating $[M - H]^-$ and $[M - 2H + Na]^-$ ions at 547.1 and 569.1 *m*/*z*, respectively, which is consistent with an isomer of TDP-D-fucose (Figure 2-18, A). However, the assignment of **Y** as TDP-D-quinovose (**44**) was ruled out due to the different HPLC retention time and relative inertness to DesII catalysis. Species **Y** was relatively stable and collected in quantities suitable for ^{1}H NMR analysis. All coupling constants between protons on the hexose ring are relatively small (< 6 Hz) indicating an absence of *trans*-diaxial C–H bonds (see Chapter 2.2.5 and Figure 2-7). Thus, species **Y** seemed to have a diaxial configuration of the vicinal hydroxyl groups at C3 and C4, and it was assigned as TDP-6-deoxy-D-gulose (**98**), the C3 epimer of TDP-D-fucose. Despite the stability of TDP-6-deoxy-D-gulose (**98**) versus the 3-keto products, the ^{1}H NMR spectra nevertheless demonstrated significant contamination due to its decomposition to a 1,2-cyclic sugar-phosphate and thymidine monophosphate during workup. Therefore, a TDP-6-deoxy-D-gulose (**98**) standard was prepared from an incubation of TDP-4-keto-6-deoxy-D-glucose (**39**) with FdtA followed by reduction with $NaBH_4$ (see Chapter 2.2.3.3, Figure 2-3). This yielded a mixture of **86** and **98**, which was used to confirm the assignment of **Y** as **98** by both HPLC coinjection and ^{1}H NMR spectroscopy.

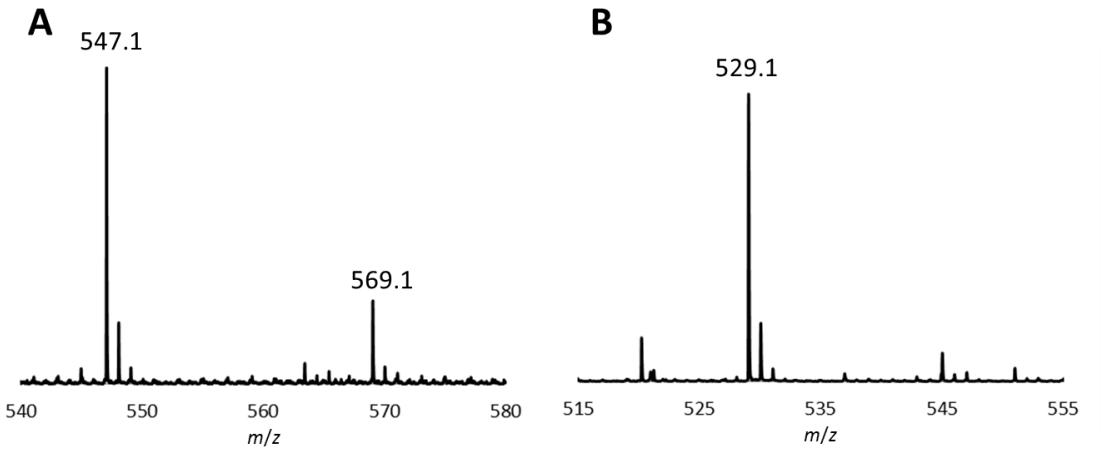


Figure 2-18. ESI-MS of A: isolated HPLC peak **Y**, and B: peak **Z** in negative ion mode.

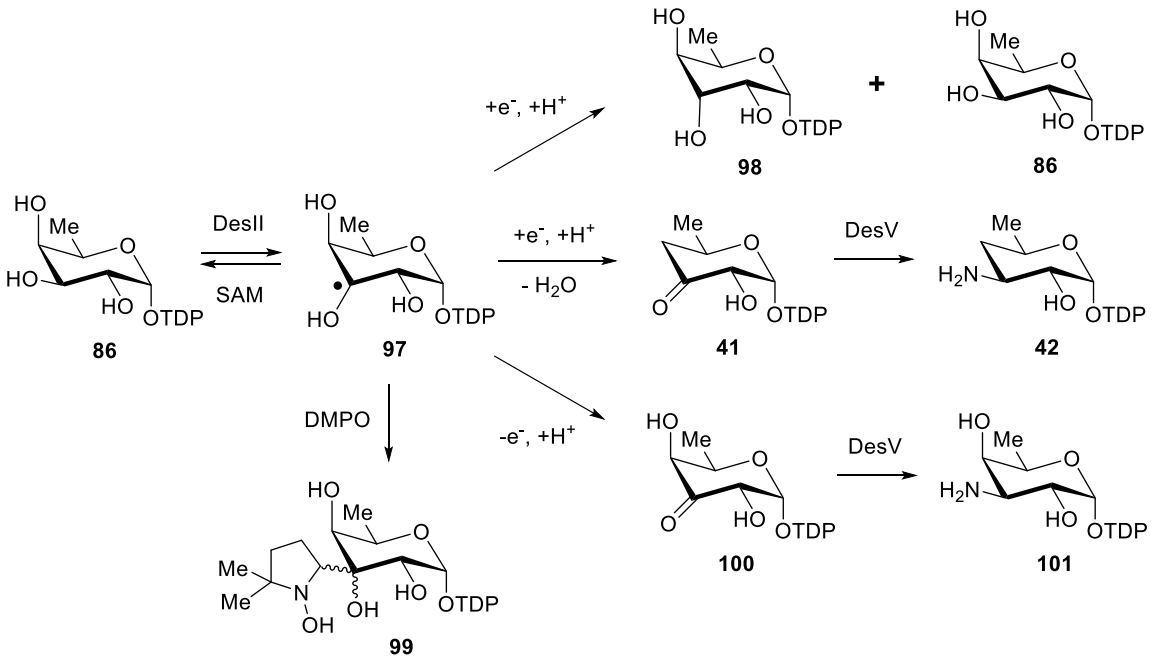


Figure 2-19. Reactions catalyzed by DesII when TDP-D-fucose (**86**) serves as the substrate.

The formation of TDP-6-deoxy-D-gulose (**98**) from TDP-D-fucose (**86**) implies that the C3 radical intermediate (**97**, Figure 2-19) is reduced via a net H-atom transfer. In buffer containing at least 95% deuterium, an approximately four-fold incorporation of deuterium versus protium into the product **98** was observed by ESI-MS. In addition, **98**, the major distinguishable product (~80%) in H_2O, was the minor product in D_2O (< 30%) indicating a solvent deuterium kinetic isotope effect on the partitioning of **97** between the different routes of decomposition. To determine the stereospecificity of net H-atom transfer to **97**, the deuterium enrichment of the residual TDP-D-fucose in the D_2O buffers was analyzed (see Chapter 2.2.6). Despite being complicated by the isotope effect disfavoring net H-atom transfer to **97**, small but measurable increases in the deuteration of the residual substrate (**86**) were detected by mass spectrometry. The regeneration of **86** suggested that net H-atom transfer is possible to both faces of the C3 radical. Finally, reactions run in the presence of the radical trap 5,5-dimethyl-1-pyrroline-*N*-oxide (DMPO) decreased production of all decomposition products, but not consumption of **86**, and resulted in two new species observable by HPLC with retention times of 17 min and 23 min (Figure 2-20). ESI-MS analysis of these collected peaks in both cases demonstrated signals at 660.1 *m*/*z* (negative ion) as well as 684.2 and 706.1 *m*/*z* (positive ions), consistent with $[M - H]^-$, $[M + Na]^+$ and $[M - H + 2Na]^+$ ions, respectively, of a DMPO adduct **99**, which would have a neutral monoisotopic mass of 661.16 Da. These results implicate that the enzyme–intermediate complex is destabilized by the axial substituent at C4 of **86**, permitting access to the substrate radical **97** by solvent, and that the net H-atom (or electron coupled proton) transferred to C3 originates from a solvent-exchangeable source.

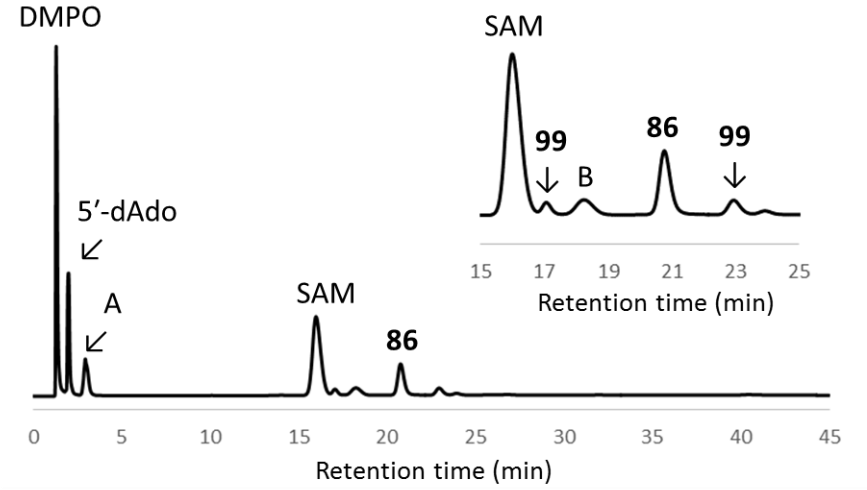


Figure 2-20. HPLC trace of DesII reaction with TDP-D-fucose (**86**) in the presence of 100 mM DMPO. Two new products corresponding to the DMPO adducts **99** are observed as shown in the inset.

The species **Z** with a retention time of 34.3 min coeluted with the deamination product **41**, and negative ion mode ESI-MS of the collected peak exhibited a clear ion signal at 529.1 *m*/*z* (see Figure 2-18, B). The identity of peak **Z** as the dehydration product **41** was further verified using a coupling reaction with DesV, which catalyzes the reductive amination of the C3 keto functionality of **41** in the presence of glutamate during the biosynthesis of TDP-desosamine.[36] The DesII/DesV coupled reaction with TDP-D-fucose (**86**) demonstrated the formation of two products with retention times of approximately 5.3 and 5.8 min, respectively, concomitant with the consumption of peak **Z** (Figure 2-21, Trace 1). The first HPLC peak of the two products coeluted with TDP-3-amino-3,4,6-trideoxy-D-glucose (**42**) produced in DesII/DesV reaction with TDP-4-amino-4,6-dideoxy-D-glucose (**40**) (Figure 2-21, Trace 2). These observations indicated that DesII is indeed

capable of operating as a dehydratase when the stereochemistry at C4 of TDP-D-quinovose (**44**) is inverted.

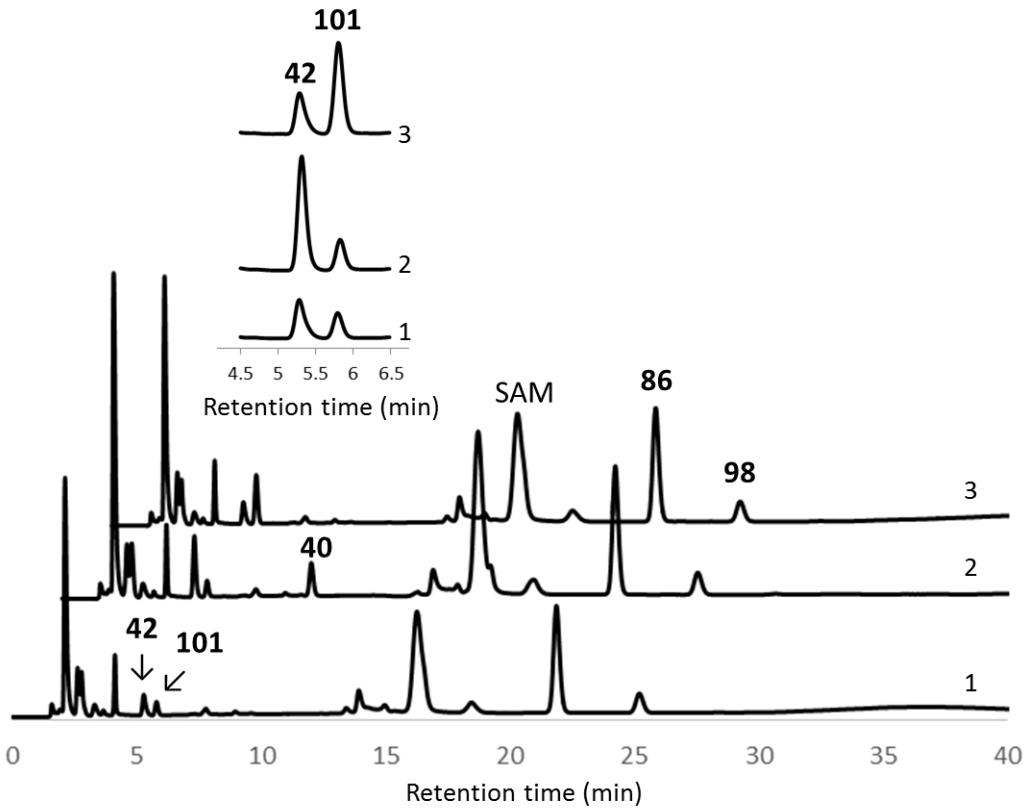


Figure 2-21. HPLC traces of coupled DesII/DesV reactions with TDP-D-fucose (**86**) along with coinjections with standard C3-amino TDP-sugars. Trace 1: DesII/DesV/TDP-D-fucose reaction; Trace 2: same as Trace 1 coinjected with a reaction of DesII/DesV/TDP-4-amino-4,6-dideoxy-D-glucose (**40**); Trace 3: same as Trace 1 coinjected with TDP-3-amino-3-deoxy-D-fucose (**101**). Retention times in Traces 2 and 3 are shifted by 2 and 4 min, respectively. The inset shows an expanded, unshifted view of the 4.5–6.5 min range where **42** and **101** elute.

The second product from DesII/DesV coupled reaction system was formed in a roughly equal quantity to **42**, and its retention time was also consistent with a TDP-aminosugar. Negative ion ESI-MS analysis of the second peak showed a signal at 546.1 *m*/*z* suggesting the $[M - H]^-$ ion of TDP-3-amino-3-deoxy-D-fucose (**101**). To confirm this assignment, a standard of compound **101** was prepared from TDP-6-deoxy-4-keto-D-glucose (**39**) using the 3,4-ketoisomerase FdtA from *Aneurinibacillus thermoaerophilus* in combination with DesV.[90,100] This standard **101** indeed coeluted by HPLC with the second DesII/DesV product from the TDP-D-fucose reaction (see Figure 2-21, Trace 3). The formation of both **42** and **101** in an approximately 1:1 ratio in the DesII/DesV reaction implies that the substrate radical **97** partitions between dehydration and dehydrogenation in a roughly equal extent without racemization at C4. It should be noted that the dehydrogenation product TDP-3-keto-D-fucose (**100**) was not directly observed during the reaction of DesII alone presumably due to the poor stability of 3-keto TDP-sugars.[92]

To summarize, TDP-D-fucose (**86**) can be recognized as a substrate for DesII despite having the inverted stereochemistry at C4 as compared to the equatorial hydroxyl group at C4 of TDP-quinovose (**44**). The majority of substrate radical **97** undergoes epimerization at C3 to generate TDP-6-deoxy-D-gulose (**98**) and likely the regeneration of **86** via net H-atom return from a solvent-exchangeable source. The remainder of **97** results in an approximately equal competition between elimination and dehydrogenation pathways (see Figure 2-19). These results imply that the partitioning of the substrate radical leading to different reaction outcomes is determined in at least in part by the stereochemical configuration of the radical intermediate.

2.3.3 DesII Reaction with TDP-6-deoxy-D-mannose (87)

The fact that DesII does not catalyze the elimination of the hydroxyl at C2 in either TDP-4-amino-4,6-dideoxy-D-glucose (**40**) or TDP-D-quinovose (**44**) is hypothesized to be due to the orthogonal orientation of the C2 substituent in the enzyme active site as observed by EPR characterization of the radical intermediate generated from **44**.[42] This hypothesis suggests that dehydration of the C2 center may take place if overlap between the C2–OH σ-bond and the *p*-orbital at C3 is increased by a conformational change in the radical intermediate. In this respect, TDP-6-deoxy-D-mannose (**87**), the C2 epimer of **44** in which the C2 hydroxyl group is in axial position, was tested as a potential substrate for DesII. Similar to TDP-4-amino-4-deoxy-D-fucose (**85**), a deuterium was introduced at C3 of **87** (**87D**, see Figure 2-24) as a probe to monitor the reaction progress.

The consumption of SAM and the production of 5′-dAdo were observed upon incubation of DesII with **87D** and SAM in the presence of $Na_2S_2O_4$ (Figure 2-22). It appeared that SAM was consumed more rapidly compared to **87D** in the presence of DesII as estimated from changes in the HPLC peak intensities. Accordingly, 5′-dAdo generated from DesII catalysis over a 6 h incubation was not completely deuterated (see Figure 2-23, A and B). The deuterium content of 5′-dAdo measured by ESI-MS was 80% and 54% in two individual trials, which suggested the uncoupled cleavage of SAM. The uncoupled reductive homolysis of SAM is commonly observed with radical SAM enzymes in the absence of substrate.[3,101] In both trials, the residual **87D** after reaction remained deuterated (ca. >95%, Figure 2-23, C and D; calculated for $C_{16}H_{24}DO_{15}N_2P_2^-$ $[M - H]^-$: 548.0798 *m*/*z*, found: 548.1 *m*/*z*; calculated for $C_{16}H_{23}DO_{15}N_2P_2Na^-$ $[M - 2H + Na]^-$: 570.0618 *m*/*z*, found: 570.1 *m*/*z*). The formation of 5′-deoxyadenosyl-5′-sulfinic acid (**89**) was also observed.

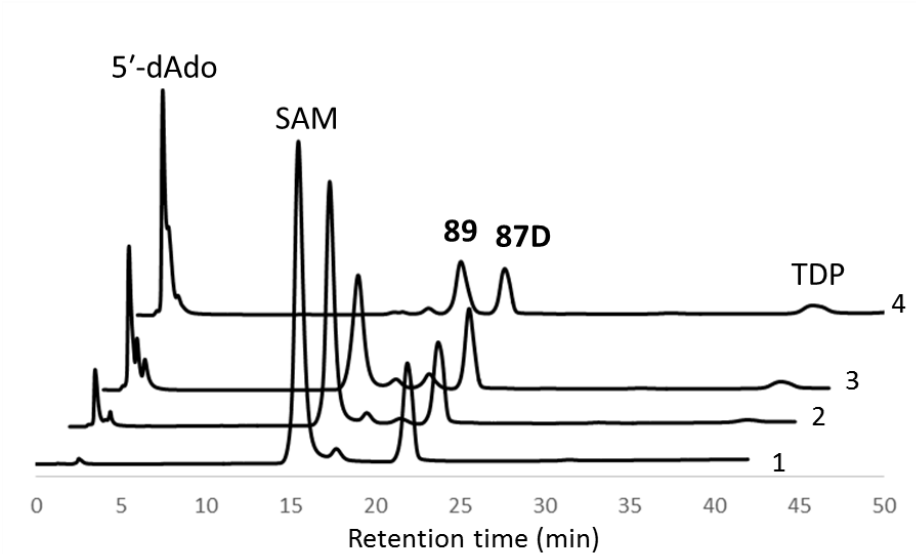


Figure 2-22. HPLC traces of DesII reaction with TDP-[3-^{2}H]-6-deoxy-D-mannose (**87D**). Trace 1 was measured after 30 min without DesII. Traces 2, 3, and 4 were measured after 30 min, 4 h, and 22 h in the presence of DesII, respectively. The retention times in traces 2, 3, and 4 are shifted by 2, 4, and 6 min for presentation purposes. The consumption of **87D** and SAM along with the production of 5′-dAdo and TDP was observed. The production of 5′-deoxyadenosyl-5′-sulfinic acid (**89**) was also noted.

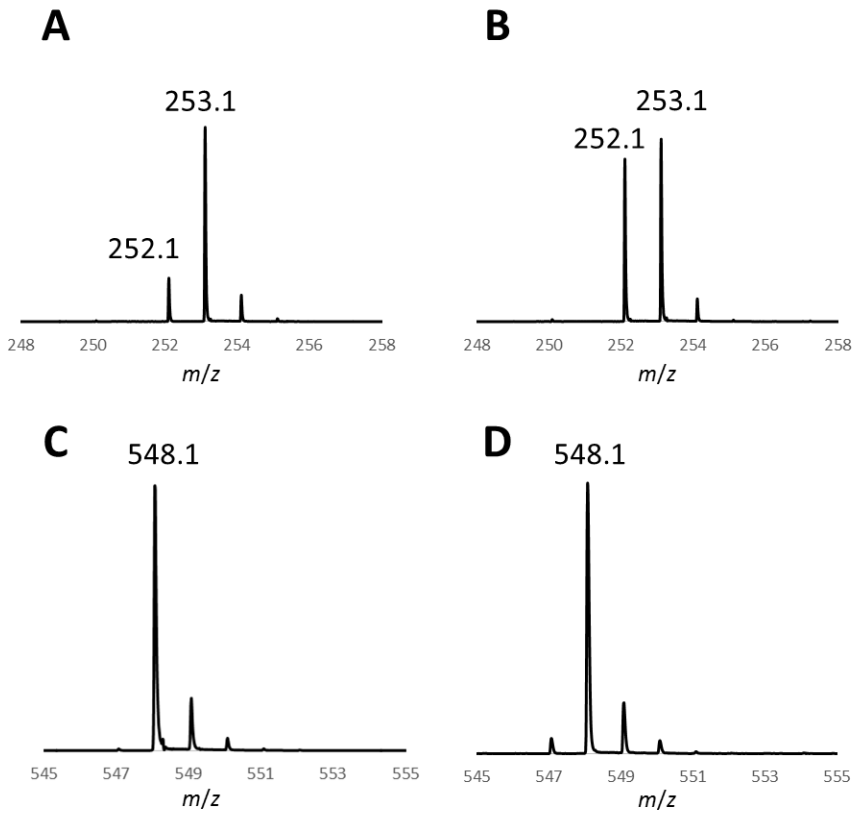


Figure 2-23. ESI-MS of A, B: 5′-dAdo generated from the DesII reaction with **87D** in two individual trials (positive ion mode), C, D: residual **87D** after 6 h in two individual trials (negative ion mode).

While there was no observable TDP-sugar products, TDP was generated during the catalysis (calculated for $C_{10}H_{15}O_{11}N_2P_2^-$ [M – H]$^-$: 401.0157 m/z, found: 401.0 m/z). TDP was not formed without DesII under the same reaction condition. The fractional conversion of **87D** to TDP estimated by the HPLC peak intensity was approximately 20% after 6 h, which is considered significant compared to less than 5% production from the DesII reaction of TDP-D-quinovose (**44**). It was previously reported that a TDP sugar with a C3 ketone functionality decomposes to yield the dihydropyran likely by deprotonation at C2 followed by elimination of TDP.[92] Thus, the formation of TDP from **87D** by DesII

suggested that a TDP-3-keto-sugar may be produced during the catalysis by DesII, and readily decomposes to generate a dihydropyran and TDP. The possible immediate products from the DesII reaction with **87D** and their degradation products are shown in Figure 2-24. The dehydrogenation reaction that is similar to the reaction of TDP-D-quinovose with DesII would produce TDP-3-keto-6-deoxy-D-mannose (**102**) that is degraded to (2*R*,3*R*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran (**105**) by eliminating TDP. In contrast, dehydration at C2 would initially produce TDP-3-keto-2,6-dideoxy-D-glucose (**103**) and then proceed to degradation to form (2*R*,3*R*)-2-methyl-3-hydroxy-4-keto-2,3-dihydropyran (**106**). Another possibility would be the formation of the C4 dehydration product **104**; however, it is less likely considering that the dehydration at C4 of TDP-D-quinovose (**44**) is not observed.[40,41]

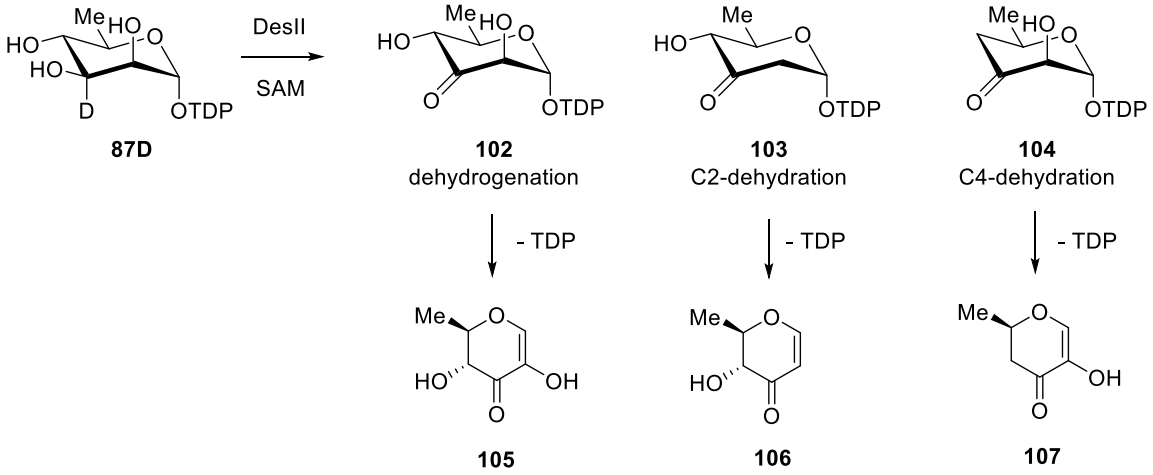


Figure 2-24. Possible reaction products from the DesII reaction with **87D**.

In an effort to characterize the TDP sugar product in the reaction between DesII and **87D**, a DesV coupled assay was performed. This resulted in only a tiny peak observed by HPLC that may correspond to a TDP-amino sugar; however, ESI-MS of this new compound did not yield any detectable signal due to low concentration. The limited conversion and rapid degradation also made the attempt to reduce the 3-keto sugar product by $NaBH_4$ unsuccessful. Instead, an effort was made to characterize the degradation product(s). A standard of dihydropyran **105** was enzymatically prepared from TDP-D-glucose (**38**) using RfbB and Tyl1a according to the published procedure.[92] A standard of compound **106** was chemically synthesized.

Standards **105** and **106** eluted very early with the retention times of less than 2 min when analyzed by anion exchange HPLC,[92] which complicated the ESI-MS analysis due to many other components in the reaction mixture that elute in this region. HPLC analysis using a C18 column resulted in longer retention times and thus better resolution of **105** and **106** (see Figure 2-25). When the reaction mixture with DesII and **87D** was analyzed under the same condition using C18 column, however, there was no peak having the same retention time as **105** or **106**. Guided by the elution profile of standard compounds, the elution fraction expected to contain the degradation product(s) was collected and analyzed by ESI-MS in positive ion mode. Neither **105** nor **106** was detected. The major species found in the reaction mixture was adenine (calculated for $C_5H_6N_5^+$ $[M + H]^+$: 136.0623 *m*/*z*, found: 136.1 *m*/*z*), which was observed by HPLC with a retention time of 14.7 min. The characterization of degradation products **105** and **106** by ESI-MS turned out to be challenging due to their intrinsic weak ionization, which was further compounded by the presence of contaminant(s). The identity of the degradation product is currently being investigated.

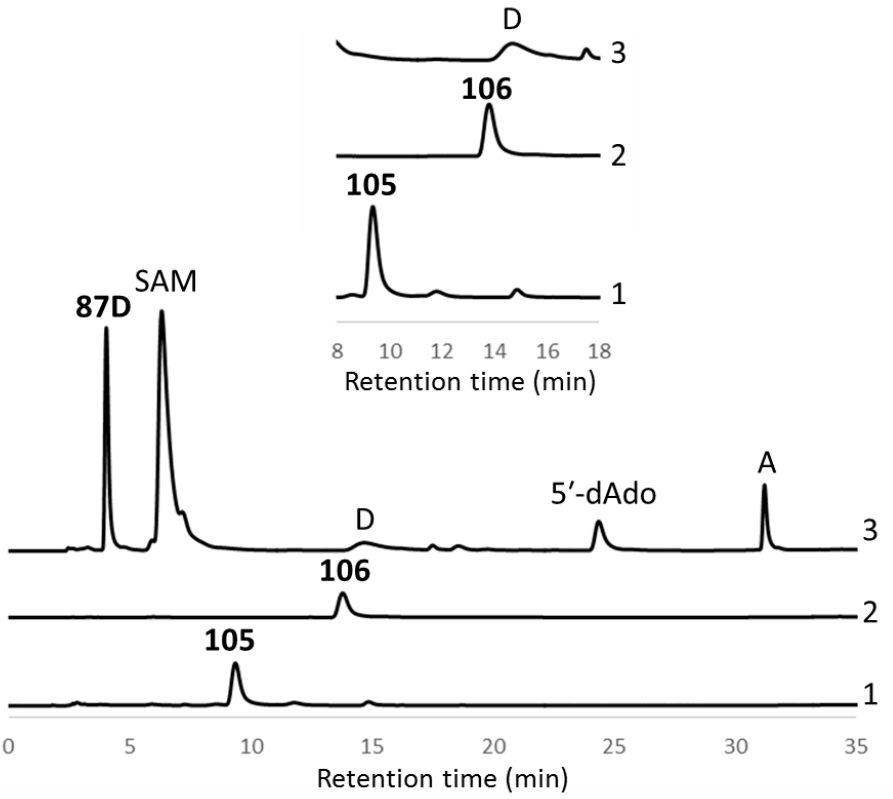


Figure 2-25. HPLC trace of DesII reaction with **87D** analyzed using C18 HPLC. Traces 1 and 2 show the injection of the standard compounds **105** and **106**, respectively. The DesII reaction with **87D** in the presence of SAM and $Na_2S_2O_4$ after 4 h is shown in trace 3. The production of 5′-dAdo and MTA (peak A) was noted. The broad peak eluting at 14.7 min (peak D) was found to be adenine. The traces from 8 min to 18 min are shown enlarged in the inset.

2.3.4 Mechanistic Implications

While structural information on the active site of DesII is not available yet, the dual activities of DesII capable of catalyzing both deamination and dehydrogenation may be explained at least in part by the geometric configuration of the substrate in the active site. In the case of TDP-D-quinovose (**44**), its C3-centered substrate radical is assumed to bind in a chair-like conformation, which results in poor hyperconjugation between the *p*-orbital at the radical center and the adjacent C–O bonds at both C2 and C4. Stalling of the catalytic cycle for the elimination reaction allows an active site base residue to deprotonate the C3 hydroxyl group facilitating the oxidation of radical intermediate by electron transfer back to the [4Fe-4S] cluster. In contrast, it is hypothesized that the radical intermediate derived from the biosynthetically relevant substrate TDP-4-amino-4,6-dideoxy-D-glucose (**40**) adopts a different binding configuration due to the protonated amino group at C4. If correct, the periplanar configuration between the *p*-orbital and the C4 amino group would account for the observed lyase activity of DesII.

The deamination reaction of TDP-4-amino-4-deoxy-D-fucose (**85**) catalyzed by DesII supports the current working model that the partitioning of radical intermediates depends at least in part on the binding configuration. Despite the inverted stereochemistry at C4, the presence of protonated amino group may reinforce the alternative binding configuration that optimizes the orbital overlap required for elimination reaction, while maintaining the orthogonal orientation with respect to the C2 substituent. There was no evidence for dehydrogenation taking place, suggesting that partitioning into the deamination pathway is significantly greater than oxidation of the radical intermediate. Moreover, the fact that the deuterium at C3 of **85D** is not exchanged implies that forward partitioning to elimination of the C4 amino group is fast compared to either the reverse reaction or quenching of the substrate radical by a solvent-accessible source. The course of

the DesII reaction with **85** contrasts with the reaction of C3-fluorinated substrate analogs (*i.e.*, TDP-[3-^{2}H]-4-amino-3,4,6-trideoxy-3-fluoro-D-glucose and TDP-[3-^{2}H]-3,6-dideoxy-3-fluoro-D-glucose) employed in a separate study. It was shown for C3-fluorinated analogs that the hydrogen atom abstraction by 5′-dAdo to form the substrate radical intermediate is reversible, and that the reverse reaction to regenerate SAM and substrate dominates when forward partitioning is impeded.[102]

The DesII reaction with TDP-D-fucose (**86**) demonstrates the diverse reaction pathways available to a radical intermediate. Upon inversion of stereochemistry at C4, the binding interaction between the enzyme and the radical intermediate from **86** is presumed to be destabilized, predominantly leading to the reduction of substrate radical via net H-atom transfer. In addition, a roughly equal partitioning between dehydration and dehydrogenation without racemization at C4 was observed. The dehydratase activity of DesII with **86** implies that perturbations of the stereochemistry can induce C–O bond cleavage, arguing against the hypothesis that dehydration is slow due entirely to the poor leaving group ability of hydroxyl group. These results support the mechanistic hypothesis in which the fate of the radical intermediate is governed at least in part by stereochemical configuration.

The reaction outcome of incubation of DesII and TDP-6-deoxy-D-mannose (**87**) is of particular interest due to the inversion of stereochemistry at C2 position. While **87** can be processed by DesII as a substrate, the key question whether or not the C2 hydroxyl group is eliminated has not been answered yet. The TDP sugar product formed from the reaction appears to be unstable because of the C3 ketone functionality in accord with the previous report.[92] Nevertheless, the formation of 5′-dAdo with the deuterium incorporated indicates the relaxed substrate specificity of DesII that can recognize the alternative substrate with the C2 axial substituent. The reverse reaction to regenerate the substrate

from the substrate radical seems to be hindered as reflected by the retention of deuterium on the residual substrate **87D**. Further experiments are required to fully characterize the degradation product, from which the structure of the initial TDP sugar product can be deduced.

2.4 CONCLUSION

DesII is the radical SAM enzyme that catalyzes the deamination reaction leading to C4 deoxygenation during the biosynthesis of TDP-desosamine (**43**) in *S. venezuelae*. While it operates as a lyase that eliminates the C4 amino group of TDP-4,6-dideoxy-3-keto-D-glucose (**40**) in the desosamine biosynthetic pathway, DesII can catalyze an oxidation reaction when TDP-D-quinovose (**44**) is provided as the substrate. The principle underlying discrimination between the deamination and dehydrogenation pathways has been studied by employing several substrate analogs with inverted stereochemistry at C2 or C4. DesII catalyzes the deamination of TDP-4-amino-4-deoxy-D-fucose (**85**), the C4 epimer of **40** which bears an axial amino substituent. This is consistent with the current mechanistic hypothesis for deamination and the proton transfer steps it entails. In the DesII reaction with TDP-D-fucose (**86**) as a substrate, the majority of the reaction flux results in epimerization at C3. However, the remainder leads to both dehydration and dehydrogenation products to an approximately equal extent, indicating that stereochemical perturbations in the substrate allow partitioning into the elimination pathway. TDP-6-deoxy-mannose bearing a C2 axial hydroxyl group (**87**) can also be recognized by DesII as a substrate, although the initial TDP-sugar product appears to be readily degraded. The identity of the degradation product is currently being investigated. Taken together, it

appears that DesII exhibits a relaxed substrate specificity regarding the stereochemistry at C2 and C4 positions despite lower catalytic efficiency. These results are consistent with the hypothesis that substrate geometry in the active site plays an important role in controlling the fate of radical intermediates during the catalytic cycle of DesII.

Chapter 3: Biosynthetic Studies of Formycin A and Coformycin

3.1 INTRODUCTION

Formycin A (**51**) is a *C*-nucleoside analog of adenosine (**53**), isolated from *Nocardia interforma* and *Streptomyces kaniharaensis* SF-557 (Figure 3-1).[51,52] Formycin A is characterized by the *C*-glycosidic linkage between ribose and the unique pyrazolopyrimidine nucleobase, compared to its *N*-nucleoside counterpart adenosine with an imidazolopyrimidine nucleobase. The formycin-producing strains also produce another nucleoside metabolite, coformycin (**108**), which possesses an unusual 1,3-diazepine ring.[52,103] Although coformycin alone has no antibacterial activity, the presence of coformycin potentiates the activity of formycin A by inhibiting the adenosine deaminase that converts formycin A into the inactive analog formycin B (**52**).[104] A similar synergetic relationship is found between 2′-deoxycoformycin (**109**, pentostatin or Nipent™) and vidarabine (**110**, 9-*β*-D-arabinofuranosyladenine), isolated together from *Streptomyces antibioticus*.[105] While 2′-deoxycoformycin (**109**) is an anticancer drug used for treatment of hairy cell leukemia,[106] vidarabine (**110**) is an antiviral nucleoside active against herpes simplex virus.[107] The D-ribose of adenosine (**53**) is replaced with D-arabinose in **110**, which results in competitive incorporation of **110** into DNA to interfere with viral DNA synthesis.[108] The antiviral activity of **110** is enhanced by **109** due to the inhibitory effect of **109** on adenosine deaminase.[108,109]

The part of this chapter has been submitted for publication and under review. I was the primary author (Ko, Y.; Wang, S.-A.; Ogasawara, Y.; Ruszczycky, M. W.; Liu, H.-w. Identification and characterization of enzymes catalyzing pyrazolopyrimidine formation in the biosynthesis of formycin A. *submitted*).

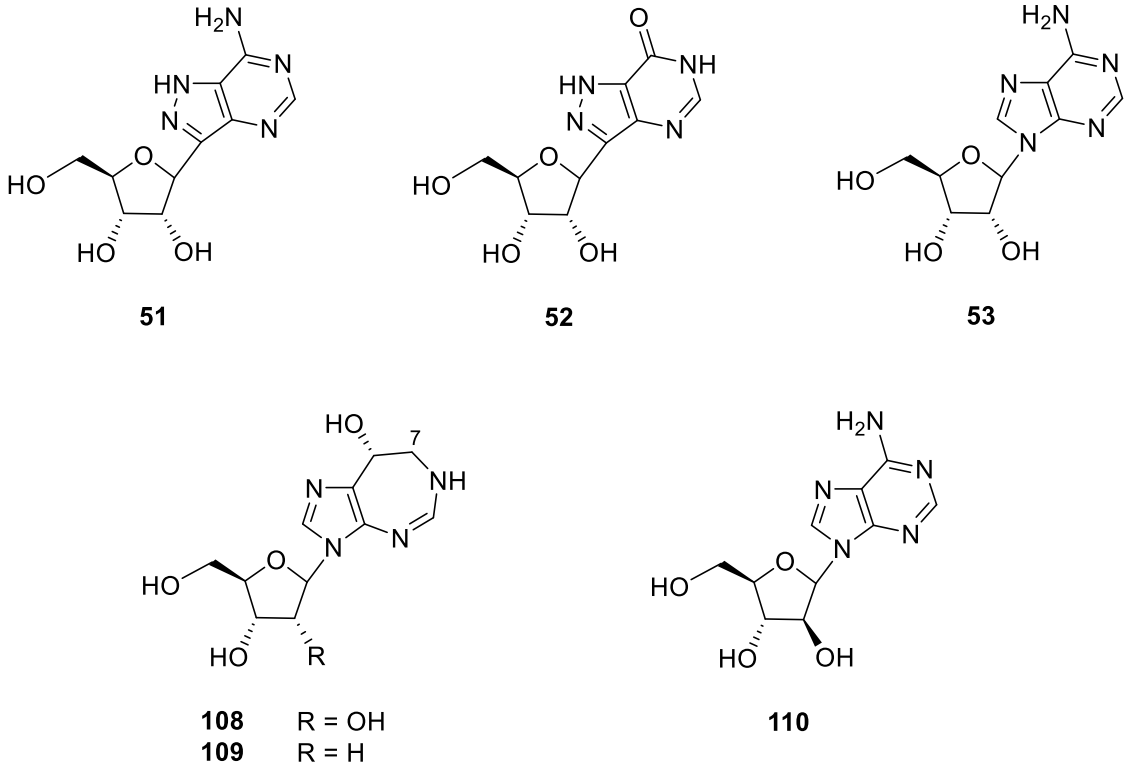


Figure 3-1. Structures of formycin A (**51**) and several related compounds: formycin B (**52**), adenosine (**53**), coformycin (**108**), 2′-deoxycoformycin (**109**), and vidarabine (**110**).

Various enzymes that recognize adenosine (**53**) as a substrate are inhibited by formycin A (**51**) due to their close structural similarity. Formycin A and its analogs have been shown to be a potent inhibitors of bacterial purine nucleoside phosphorylase, whereas they are inactive against mammalian enzymes.[110,111] Likewise, formycin A inhibits adenosine kinase from *Mycobacterium tuberculosis*, which is involved in the purine salvage pathway,[112] and *E. coli* 5′-methylthioadenosine (MTA)/*S*-adenosylhomocysteine (SAH) nucleosidase, which is involved in the catabolism of MTA and SAH.[113] Therefore, formycin A is of interest for its potential in selective drug design. Furthermore, it has been reported that formycin A exhibits antiproliferative properties with respect to various cancer

cells[114,115] and antiviral activity against influenza virus A14[116] and human immunodeficiency virus type 1.[117]

Although extensive research has been devoted to study the biological activities of formycins, investigation of their biosynthesis has been rare and slow. Early *in vivo* and *in vitro* studies have shown that the pyrazolopyrimidine ring of formycin A is derived from L-glutamate (**75**) and L-lysine (**76**),[68-70] and formycin B 5′-phosphate (**77**) is a likely intermediate in the biosynthetic pathway of formycin A (see discussion in Chapter 1.2.3).[66] However, the biosynthetic gene cluster and enzymes responsible for such transformations between these precursors and the final product have not yet been identified.

The biosynthesis of coformycin (**108**) is presumed to be associated with that of formycin A given their similar production profiles from the same bacterial strain and their synergetic biological activities. The diazepine ring found in coformycin (**108**) and its derivatives rarely occurs in nature. *In vivo* isotope labeling experiments in *S. antibioticus* indicated that the methylene group at C7 of 2′-deoxycoformycin (**109**) is derived from the C1 of ribose, and the possibility that inosine (**54**) is a direct precursor of **109** has been ruled out.[118,119] In addition, cell-free extracts of *S. antibioticus* were shown to be capable of catalyzing the reduction of 8-ketocoformycin (**119**) and 8-ketodeoxycoformycin (**120**) in the presence of NADPH. Taken together, the above observations led to the proposed biosynthetic pathway of coformycin (**108**) and 2′-deoxycoformycin (**109**, see Figure 3-2).[120] The early steps of the proposed pathway are similar to the biosynthesis of L-histidine (**111**, **112** → **116**),[121] which begin with the coupling of PRPP (**74**) and adenosine 5′-triphosphate (**111**, ATP), or 2′-deoxyadenosine-5′-triphosphate (**112**, dATP) in the case of **109**. Dephosphorylation followed by ring cleavage generates the intermediate **115**. An Amadori-type rearrangement of **115** leads to the formation of intermediate **116**, and then the 1,3-diazepine ring is formed via an aldol-like mechanism. Elimination of four-carbon

chain, dephosphorylation, and reduction of the ketone group produces **108** or **109**. Although the non-phosphorylated 8-keto compounds (**119** and **120**) were identified as the biosynthetically relevant intermediates, whether the dephosphorylation is followed by the reduction, or vice versa, is not known.

Chapter 3 describes biosynthetic studies of formycin A (**51**) and coformycin (**108**). In order to discover the biosynthetic gene cluster for formycin A (**51**), the genome of *S. kaniharaensis* was sequenced. The preliminary genomic analysis revealed two sets of *pur* homologous genes, which are involved in the later stage of the biosynthesis of adenosine 5′-monophosphate (**127**, AMP). This finding led to the hypothesis that the assembly of the pyrimidine ring in both formycin A and AMP follows a similar pathway. A cosmid library was constructed and screened for the second copy of *pur* genes according to the hypothesis. As a result, one cosmid was identified to harbor the second set of *pur* genes and likely the entire gene cluster of formycin A. Bioinformatics analysis permitted assignment of putative functions to each gene, from which a biosynthetic pathway for formycin A could be proposed. The conversion of carboxyaminopyrazole ribonucleotide (**128**) to formycin A (**51**) was demonstrated by incubation using Pur-like enzymes encoded in the gene cluster. Moreover, one of the genes flanked by the end of the putative gene cluster of formycin A was shown to encode an NADPH-dependent reductase that catalyzes the last step of the biosynthesis of coformycin. Ultimately, this research aims to elucidate the complete biosynthetic pathway of formycin A and coformycin, and examine the detailed reaction mechanisms of the enzymes involved.

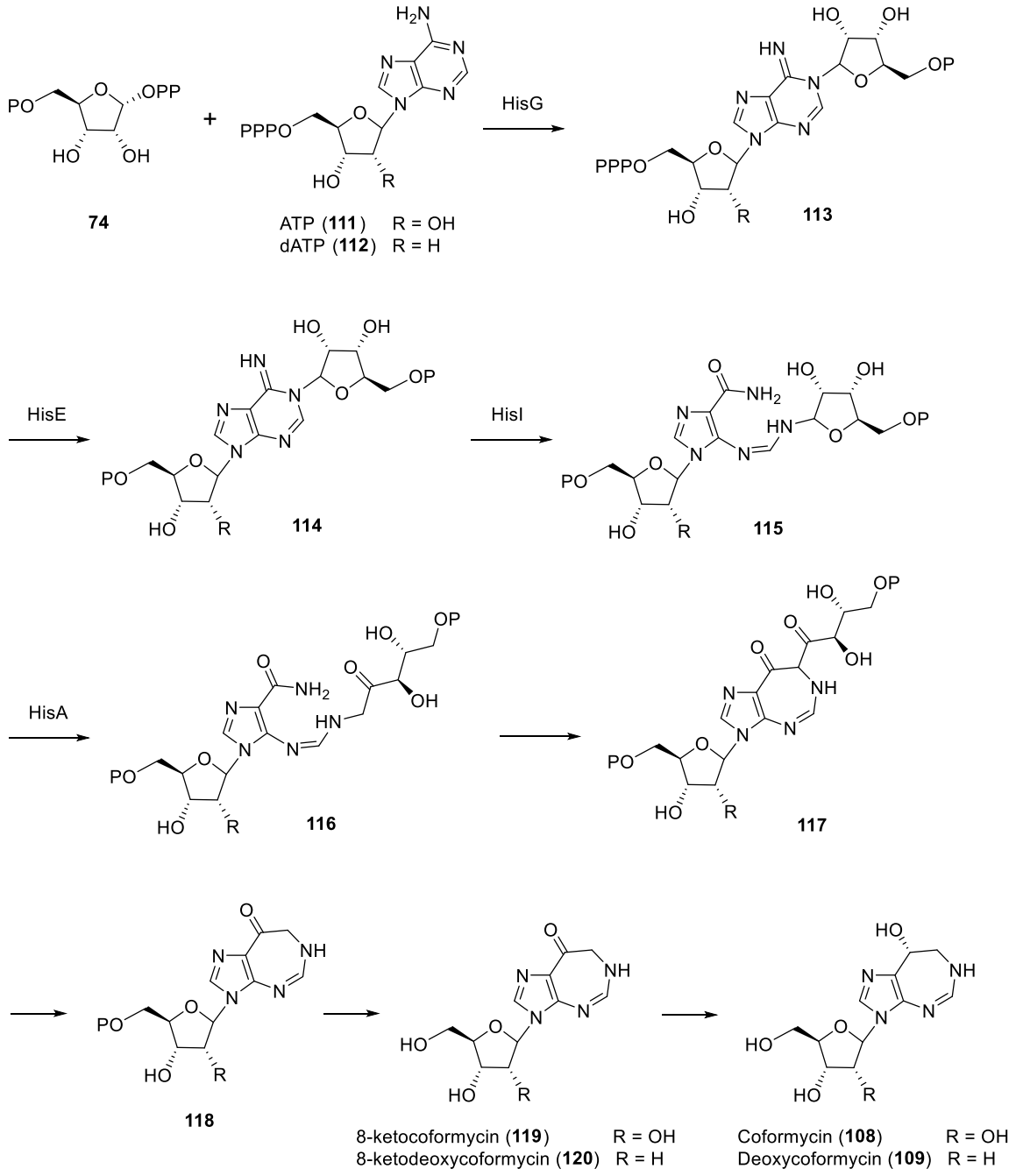


Figure 3-2. Proposed biosynthesis of coformycin and deoxycoformycin.

3.2 EXPERIMENTAL PROCEDURES

3.2.1 General

Kits for isolation of genomic DNA and plasmid DNA were purchased from Life Technologies (Carlsbad, CA) and Qiagen (Hilden, Germany), respectively. A kit for construction of cosmid libraries was purchased from Agilent Technologies (Santa Clara, CA). Formycin A (**51**) was purchased from Berry & Associates, Inc (Dexter, MI). All chemicals and reagents were purchased from Fisher Scientific (Pittsburgh, PA) or Sigma-Aldrich (St. Louis, MO), and were used without further purification unless otherwise specified.

The bacterial strains *Streptomyces kaniharaensis* (ATCC21070) and *Streptomyces antibioticus* (NRRL 3238) were obtained from the American Type Culture Collection and the Agricultural Research Service Culture Collection, respectively. *E. coli* DH5α and *E. coli* BL21 Star (DE3) from Invitrogen (Carlsbad, CA) were used as the host for routine cloning and heterologous protein expression.

The genomic DNA of *Streptomyces kaniharaensis* SF-557 was sequenced using a Roche 454 sequencer at the Genomic Sequencing and Analysis Facility (GSAF) at the University of Texas at Austin. Routine DNA sequencing was performed by the Institute for Cellular and Molecular Biology Core Facilities at the University of Texas at Austin. All high-resolution mass spectrometry (HRMS) was performed at the Mass Spectrometry Facility (MSF) of the Department of Chemistry, University of Texas at Austin. Standard genetic manipulations of *E. coli* were performed as described by Sambrook and Russell.[122]

3.2.2 Genomic Sequencing of *S. kaniharanesis*

Approximately 30 μg of genomic DNA of *S. kaniharaensis* was isolated using PureLink genomic DNA kits according to the manufacturer's protocol. The isolated genomic DNA was sequenced using a Roche 454 sequencer at the GSAF which resulted in 3011 contiguous sequences (contigs) with an average size of 2641 bp. The MIRA program was used to assemble the contigs,[123] and ORFs were defined with Glimmer using a hidden Markov model to yield 9527 open reading frames (ORFs) in total.[124] The a biological function for each ORF was proposed based on BLAST searches.

3.2.3 Construction and Screening of the Cosmid Library

The vector pOJ446 was digested with *Hpa*I and treated with calf intestinal alkaline phosphatase (CIP). After purification, pOJ446/*Hpa*I/CIP was digested with *Bam*HI. The genomic DNA of *S. kaniharaensis* was partially digested by *Sau*3AI and then purified by phenol/chloroform/isoamyl alcohol (PCI) extraction, chloroform extraction, and isopropyl alcohol precipitation. The partially digested genomic DNA was treated with CIP and purified by PCI extraction, chloroform extraction, and isopropyl alcohol precipitation. The partially digested genomic DNA was ligated with the pOJ446/*Hpa*I/CIP/*Bam*HI vector. The ligated DNA was packaged into phage particles using Gigapack III XL packaging extract and introduced into *E. coli* XL1-Blue MRF′ according to the manufacturer's protocol. The cells were plated onto Luria broth (LB) agar and selected by using apramycin (50 μg/mL). The positive colonies were collected in LB containing apramycin (50 μg/mL), incubated at 37 °C overnight, and stored at –80 °C with 20% glycerol.

For screening, the positive colonies harboring the cosmid library were grown at 37 °C overnight and DNA was obtained by miniprep. The cosmid library was screened by polymerase chain reaction (PCR). The primers used were: 5′-GCTGGCCTGATC-

CGCGAACC-3′ (forward for *purA* homolog), 5′-GCCGATGGTCCGGTCCACGC-3′ (reverse for *purA* homolog), 5′-CCAGTTCTCCGTCGCACTCG-3′ (forward for *purC* homolog), 5′-TGCAGACCGGGGTCCCTCGG-3′ (reverse for *purC* homolog), 5′-GTGCGGCGTCCACGACCAGG-3′ (forward for *purH* homolog), 5′-GCTGGAGGG-TCGCGGCCACG-3′ (reverse for *purH* homolog).

The K24C cosmid was found to contain all *purA*, *purC*, and *purH* homologs, and the K5B cosmid was found to contain the *purC* homologous gene. Thus, the K24C cosmid was subjected to sequencing by primer walking, and the results were assembled using MIRA. ORFs were defined by Glimmer and each gene was annotated based on a comparison with the non-redundant database using BLAST.

The obtained sequences of *purA*, *purB*, *purC*, and *purH* homologous genes were compared to the known sequences of adenosine 5′-monophosphate (AMP) biosynthetic genes found in *Streptomyces coelicolor A3(2)*, *Streptomyces griseus subsp. Griseus NBRC 13350*, *Streptomyces avermitilis MA-4680*, and *Streptomyces scabiei 87.22* by BLAST. The locus tags used for analysis were: SCO3629, SGR_RS16855, SAV_4547, and SCAB_RS19930 for adenylosuccinate synthetase (*purA*); SCO1254, SGR_RS31355, SAV_7077, and SCAB_RS37365 for adenylosuccinate lyase (*purB*); SCO4071, SGR_RS19105, SAV_4146, and SCAB_RS22720 for phosphoribosylaminoimidazole-succino-carboxamide synthase (*purC* or *hemH*); SCO4814, SGR_RS13410, SAV_3444, and SCAB_RS16940 for bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase (*purH*).

3.2.4 Cloning and Expression of Enzymes Encoded in the Biosynthetic Gene Cluster of Formycin A

Each gene in the putative biosynthetic gene cluster of formycin A was PCR amplified from the K24C cosmid and cloned into the pET28b(+) vector unless otherwise specified. The primers used in this study are listed in Table 3-1.

Expression of recombinant proteins was routinely performed as described below, unless otherwise specified. *E. coli* BL21 Star (DE3) cells were transformed with the resulting plasmid and grown at 37 °C in LB medium containing kanamycin (30 μg/mL). The overnight culture was used to inoculate 3 L of the same medium in a 100-fold dilution. These cultures were incubated at 37 °C until OD_{600} reached between 0.4–0.6. Protein expression was induced by the addition of isopropylthiogalactoside (IPTG) to a final concentration of 0.2 mM. After overnight incubation at 18 °C, the cells were harvested by centrifugation at 4500 × g for 8 min, resuspended in 80 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl, and disrupted by sonication. Cell debris was removed by centrifugation at 20000 × g for 30 min, and the supernatant was mixed by slow agitation with 5 mL of nickel-nitrilotriacetic acid (Ni-NTA) resin for 10 min at 4 °C. The slurry was transferred to a column and washed with 50 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 20 mM imidazole. The protein was eluted with 25 mL of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 250 mM imidazole, and dialyzed three times against 1 L of 100 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 10–20% glycerol prior to storage at –80 °C.

Table 3-1. Primers for cloning used in this study (restriction sites underlined).

Gene	Forward primer	Reverse primer
orf -2	5′-TTAATTGGCATATGGGTGAGCCT-3′	5′-AGAGAATTCAGACGTGCTGGAG-3′
orf -1	5′-TTCCTTCACATATGACTTCCCGTACTCC-3′	5′-AAAGAATTCACCCACCGGAGACC-3′
forA	5′-GCCATCCTCATATGTCCGAGCGAG-3′	5′-CGAGAATTCAGCGCGGCGG-3′
forB	5′-ATTATCCACATATGACCGGCCTGG-3′	5′-AACTCTCGAGTCAGCTCCTCTCC-3′ (for pET-28b(+) vector), 5′-AAATCTCGAGGCTCCTCTCCGC-3′ (for pET24b(+) vector)
forC	5′-ATAAAGTCCATATGCTGGCCGTACTCG-3′	5′-GCAGAATTCACGACCGGCTCC-3′
forD	5′-AATAAGTCCATATGCTGAAGGCAATCG-3′	5′-CAGGAATTCTACTTGATCTCGATCAGC-3′
forE	5′-TAAATAGCCATATGCGACCCACCG-3′	5′-GGAGAATTCAGGCCAGCAGC-3′
forG	5′-AAAAAATACATATGATCACCGAGAAGTCC-3′	5′-AAAGAATTCAGGCCCGGCC-3′
forH	5′-TTTTTCCACATATGACCGTTCCACG-3′	5′-TTCTCTCGAGTCAGTGCGATGG-3′
forI	5′-TTGAGTGACATATGAGTGAGATGC-3′	5′-AAAGAATTCAGCGCCCCGTG-3′

Table 3-1. Continued.

forJ	5′-AAAAAGAACATATGAAGGCGGCG-3′	5′-CCAGAATTCAGCCGATCTCG-3′ (for pET-28b(+) vector), 5′-AAAACTCGAGGCCGATCTCGTGC-3′ (for pET-24b(+) vector)
forK	5′- AAAAAAGGCATATGGCGAAGCGC-3′	5′-TTATCTCGAGCTACTCTCCGTCG-3′
forP	5′-TTATTCCACATATGACCCGGATCATCGTC-3′	5′-TTAGAATTCACGCCGCCACGAAG-3′
forQ	5′-ATAATAGCCATATGCGACAGGTTG-3′	5′-ATAGAATTCACCGCTGCTCC-3′
forR	5′-TACAAGAACATATGAAGACCGACATCC-3′	5′-AGAGAATTCAGAGCAGCGGG-3′
forS	5′-AATTAACGCATATGGCACTTGACCGACCC-3′	5′-GGAGAATTCAGGAGACGACGGTGG-3′
forT	5′-TAAAATCACATATGGCTACGCGGAAGG-3′	5′-ACAGAATTCAGGACACCAGGTCG-3′
forU	5′-CAAAATCTCATATGGCCCATGTGC-3′	5′-AAAGAATTCAGGCGGAGGTGG-3′
forV	5′-AATAAGAACATATGAAGCGCTGGTCC-3′	5′-GATGAATTCTAGTCGGCCGAAGC-3′
forW	5′-ATTATTAGCATATGGTTGTTGACGGC-3′	5′-TTAGAATTCATCGGCGGTCC-3′
forX	5′-AAAAACCACATATGACCGCCAAGC-3′	5′-TATGAATTCTACTGCGCGGTGG-3′

Table 3-1. Continued.

purA	5′-ATCAAAAACATATGCCGGCACTCG-3′	5′-CAGGAATTCATGAACGAGTTGATCTGG-3′
*purB**	5′-AAAAAACATATGACTGCGCCTGC-3′	5′-AAAGAATTCAGAGGATCGCCCC-3′
purC	5′-ATAATAAACATATGAGCGGATTTGTCACC-3′	5′-AGAGAATTCAGACCCAGCTG-3′
*purH***	5′-CCCAAAACCATATGCAACAACGTCG-3′ (for the wildtype PurH of *E. coli*) 5′-TATTTACGCATATGGCCGCCAAGAACC-3′ (for the truncated PurH of *E. coli*) 5′-AAAATCATCATATGTACGACAAGACC-3′ (for PurH of *S. kaniharaensis*)	5′-TTTTCTCGAGTTAATGGCGGAAGTGG-3′ (for the wildtype PurH of *E. coli*) 5′-GGGGAATTCTTAATGGCGGAAGTGG-3′ (for the truncated PurH of *E. coli*) 5′-TTAGAATTCAGTGGAAGAAGTGC-3′ (for PurH of *S. kaniharaensis*)

* PurB was PCR-amplified from *Streptomyces avermitilis*.

** PurH was PCR-amplified from both *E. coli* and *S. kaniharaensis*. The *C*-terminal domain of *E. coli* PurH was also cloned.

3.2.5 Characterization of Construction of Pyrimidine Ring in Formycin A

3.2.5.1 Preparation of enzymes and substrates

The cloning, expression, and purification of ForV, ForR, and ForC were conducted as described in Chapter 3.2.4 (Figure 3-3). For ForD, 10% glycerol was added to all buffers during purification, and 25 mM potassium phosphate buffer (pH 8.0) was used rather than Tris buffer. This is because Tris has been reported to be an inhibitor of the reaction,[125] and the activity of yeast PurA is extremely sensitive towards chloride concentration.[126]

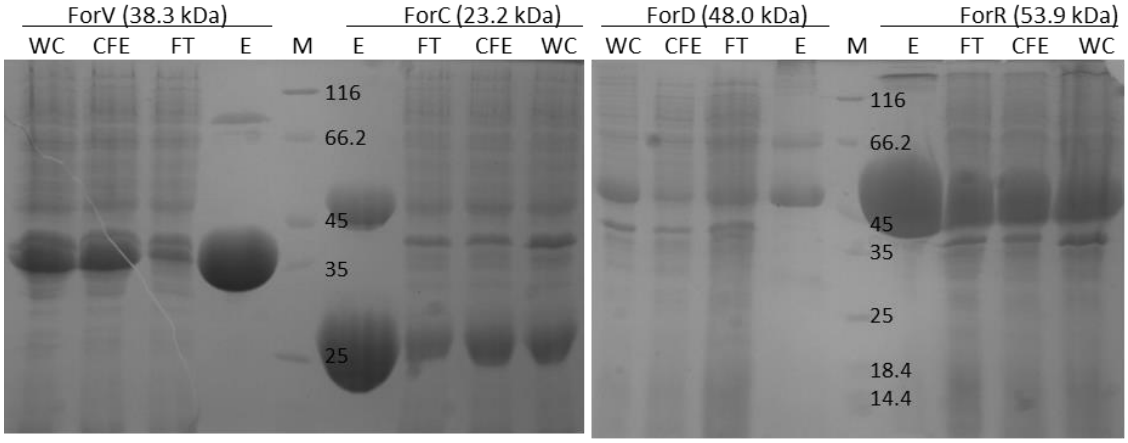


Figure 3-3. SDS-PAGE of expressed ForV, ForR, ForC, and ForD. WC: whole cell, CFE: cell-free extract, FT: flow-through, E: eluted protein, M: protein marker (Pierce unstained protein molecular weight marker).

The genuine Pur enzymes for adenosine biosynthesis were overexpressed for the purpose of activity comparison. PurC from *S. kaniharaensis* and PurH from *E. coli* and *S. kaniharaensis* were cloned and expressed as described. In addition, the truncated PurH of *E. coli* that only has the *C*-terminal AICAR formyltransferase domain was cloned and used for enzymatic preparation of intermediates. Since the *N*-terminal His_6-tagged PurA from *S.*

kaniharaensis was insoluble, it was cloned as a *C*-terminal His_6-tagged protein in the pET24b(+) vector and purified according to the procedure described for ForD. As the complete sequence of *purB* from *S. kaniharaensis* is not known, PurB from *Streptomyces avermitilis*, which has 85% identity and 92% similarity when compared to the known sequence of PurB of *S. kaniharaensis*, was expressed instead.

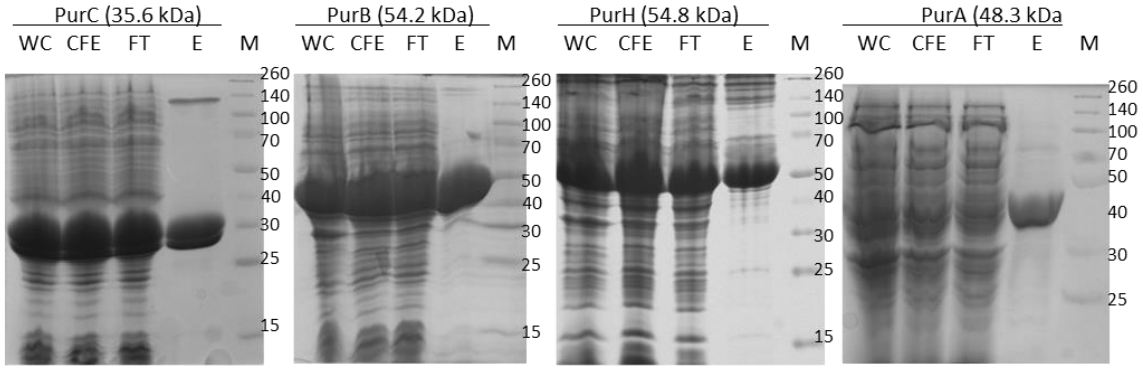


Figure 3-4. SDS-PAGE of Pur enzymes involved in adenosine biosynthesis. PurC, PurH, and PurA are from *S. kaniharaensis*, and PurB is from *S. avermitilis*. WC: whole cell, CFE: cell-free extract, FT: flow-through, E: eluted protein, M: protein marker (Spectra multicolor broad range protein ladder).

N^{10}-Formyl-tetrahydrofolate (N^{10}-formyl-THF), the donor of formyl group for PurH reaction, was anaerobically prepared according to the previously described procedure.[127] Several pathway intermediates proposed for the biosynthesis of formycin A, such as formycin B 5′-phosphate (**77**), carboxyaminopyrazole ribonucleotide (**128**), and 5-aminopyrazole-4-carboxamide ribonucleotide (**130**), were chemically synthesized by Shao-An Wang (*manuscript in preparation*).

3.2.5.2 Activity of ForV and ForR

The activity of ForV and ForR was assayed according to the procedure described for PurC with minor modifications.[128] The reaction mixture contained 0.5 mM carboxyaminopyrazole ribonucleotide (**128**), 10 mM L-aspartate, and 2 mM ATP in 50 mM Tris buffer (pH 8.0) with 20 mM potassium chloride and 6 mM magnesium chloride. After addition of 20 μM ForV and ForR, the reaction mixture was incubated at room temperature for 4 hours, and then the enzymes were removed by using YM-10 centrifugal filtration. An analytical 4 × 25 mm Dionex CarboPac PA1 anion exchange HPLC column was employed using a two-solvent system: solvent A was H_2O, and solvent B was 1.5 M ammonium acetate in H_2O (pH 7.0). The flow rate was 1 mL/min and the detector was set to 260 nm. The column was eluted with a two-phase linear gradient: 10 to 20% B in 5 min and then 20 to 30% B in 15 min. The reaction products eluted at 7.8 min and 24.7 min and were collected and lyophilized to dryness for ESI-MS analysis (negative ion mode). The peak at 7.8 min was identified to be 5-aminopyrazole-4-carboxamide ribonucleotide (**130**); calculated for $C_9H_{14}O_8N_4P^-$ [*M* – *H*]$^-$: 337.0555 *m*/*z*, found: 337.1 *m*/*z*; calculated for $C_9H_{13}O_8N_4PNa^-$ [*M* – 2H + Na]$^-$: 359.0374 *m*/*z*, found: 359.0 *m*/*z*. Peak at 24.7 min was found to be adenosine 5′-diphosphate (ADP); calculated for $C_{10}H_{14}O_{10}N_5P_2^-$ [*M* – *H*]$^-$: 426.0221 *m*/*z*, found: 426.0 *m*/*z*.

3.2.5.3 Activity of ForC

The activity of ForC was assayed following the previously described procedure for PurH.[129] The reaction mixture contained 33 mM Tris buffer (pH 7.4), 25 mM potassium chloride, 5 mM 2-mercaptoethanol, 0.5 mM 5-aminopyrazole-4-carboxamide ribonucleotide (**130**), 1 mM N^{10}-formyl-THF, 50 μM ForC, and 30 μM PurH either from *E. coli* or from *S. kaniharaensis*. After anaerobic incubation inside the glove box at room temperature, the reaction mixture was deproteinized using YM-10 centrifugal filter and

analyzed by HPLC using a method similar to that developed for the ForV/ForR reaction. The column was eluted with 10% B isocratic for 5 min, 10 to 30% B in 15 min and then 30% B isocratic for 20 min. The source of PurH did not affect the reaction profile, as the HPLC traces were identical. The reaction products eluting at 7.4 min and 14.9 min were collected and lyophilized to dryness for ESI-MS analysis (negative ion mode). The peak at 7.4 min is consistent with 5-formaminopyrazole-4-carboxamide ribonucleotide (**131**); calculated for $C_{10}H_{14}O_9N_4P^-$ $[M - H]^-$: 365.0504 *m*/*z*, found: 365.1 *m*/*z*; calculated for $C_{10}H_{13}O_9N_4PNa^-$ $[M - 2H + Na]^-$: 387.0323 *m*/*z*, found: 387.0 *m*/*z*. Peak at 14.9 min is consistent with formycin B 5′-phosphate (**77**); calculated for $C_{10}H_{12}O_8N_4P^-$ $[M - H]^-$: 347.0398 *m*/*z*, found: 347.0 *m*/*z*; calculated for $C_{10}H_{11}O_8N_4PNa^-$ $[M - 2H + Na]^-$: 369.0218 *m*/*z*, found: 369.0 *m*/*z*. The $[M - H]^-$ ion of the hydrate form was also detected at 365.1 *m*/*z*.

3.2.5.4 Activity of ForD and ForR

The ForD/ForR coupled assay was conducted based on a reported assay for PurA with minor modifications.[126] A solution of 10 μM ForD and 10 μM ForR was incubated with 1 mM formycin B 5'-phosphate (**77**), 2.5 mM L-aspartate, and 0.125 mM guanosine triphosphate (GTP) in 1 mL of 25 mM potassium phosphate buffer (pH 8) containing 16 mM magnesium acetate. To regenerate GTP, 1 mM phosphoenolpyruvate (PEP) and 2 units pyruvate kinase were also included in the reaction mixture. After incubation for 2 hours at room temperature, the enzymes were removed by filtering through a YM-10 centrifugal filter. The reaction was analyzed by HPLC using the same method as the ForV/ForR reaction.

The formation of succinylformycin 5′-phosphate (**132**) in the reaction mixture of ForD alone was not observable under the HPLC condition due to the prolonged retention of **132** on the anion exchange column. However, high resolution LC-MS analysis of the

reaction confirmed the production of **132**. HRMS (ESI, negative ion mode) calculated for $C_{14}H_{17}N_5O_{11}P^-$ [*M* – *H*]$^-$: 462.0668 *m*/*z*, found: 462.0658 *m*/*z*. The product of the ForD and ForR coupled reaction, formycin A 5′-phosphate (**78**), was eluted at 14.6 min. HRMS (ESI, negative ion mode) of **78**: calculated for $C_{10}H_{13}N_5O_7P^-$ [*M* – *H*]$^-$: 346.0558 *m*/*z*, found: 346.0556 *m*/*z*. The substrate **77** was eluted at 9.5 min under this condition.

The production of formycin A 5′-phosphate (**78**) was confirmed by treating it with calf intestinal alkaline phosphatase (CIP). Approximately 1 mM of **78** isolated from the reaction mixture of ForD and ForR was dissolved in 20 mM Tris buffer (pH 7.9) containing 50 mM potassium acetate, 10 mM magnesium acetate, and 100 μg/mL bovine serum albumin. A total of 100 units of CIP was added, and the reaction mixture was incubated at 37 °C for 1 hour. After the enzymes were removed by YM-10 centrifugal filtration, the reaction mixture was analyzed by HPLC using a Microsorb-MV 100-5 C18 column. A two-solvent system of A, 1% ammonium acetate in H_2O (pH 5.1), and B, acetonitrile, was used for elution with the flow rate set at 1 mL/min, and the UV-detector set to 260 nm. The column was eluted with 1% B for 10 min, and then a two-phase linear gradient of 1 to 10% B in 20 min and 10 to 35% B in 10 min was applied. While formycin A 5′-phosphate (**78**) was eluted at 5.4 min, formycin A (5**1**) produced from treatment with CIP was eluted at 19.1 min and co-eluted with the commercially available formycin A standard. The collected formycin A was lyophilized to dryness. HRMS (ESI, negative ion mode) calculated for $C_{10}H_{12}N_5O_4^-$ [*M* – *H*]$^-$: 266.0895 *m*/*z*, found: 266.0893 *m*/*z*.

3.2.6 Competition Experiment

The competition experiments were run by mixing both substrates, formycin B 5′-phosphate (**77**) and inosine monophosphate (**125**, IMP) in the reaction, and the ratio of **77** to **125** was monitored by measuring HPLC peak intensities as a function of fraction of reaction. Specifically, the reaction mixture contained 20 mM HEPES (pH 7.7), 1 mM magnesium chloride, 0.5 mM GTP, 5 mM L-aspartate, 1 mM PEP, 1 unit pyruvate kinase, and 3–5 μM ForD. The concentrations of formycin B 5′-phosphate (**77**) and IMP (**125**) were varied between 0.1 mM and 1 mM, while keeping the initial concentration ratio of **77** to **125** below 0.3. Also included in the reaction mixture is 0.2 mM or 0.5 mM of 5-aminoimidazole-4-carboxamide 1-*β*-D-ribofuranoside (**133**, AICAR) which was used as an internal standard. The reactions were run at room temperature, and aliquots were taken prior to the addition of enzymes and at given time points.

Each reaction mixture collected at a given time point was analyzed by HPLC using a Microsorb-MV 100-5 C18 column. The column was eluted with 1% ammonium acetate in H_2O (pH 5.1), and washed between injections with a gradient of acetonitrile up to 10%. The flow rate was 1 mL/min. The detector was set to 278 nm, which is the absorbance maxima of formycin B 5′-phosphate (**77**) ($\varepsilon_{278,77}$ = 7300 $M^{-1}cm^{-1}$), to improve the detection limit for **77** at low concentrations.[110] The relative extinction coefficient of IMP (**125**) at 278 nm was estimated to be 2292 $M^{-1}cm^{-1}$ according to Beer's Law. The retention time of **77** and **125** under this HPLC condition was 3.8 min and 4.1 min, respectively. The products, succinylformycin 5′-phosphate (**132**) and succinyladenosine 5′-monophosphate (**126**, SAMP), were eluted at 3.1 min and at 4.5 min, respectively.

The *i*-th relative substrate concentration ratio of **77** to **125** (R_i) was measured by HPLC peak intensity (h_i) and the *i*-th fraction of reaction (f_i) was determined using the

internal standard **133**, where $i = 1$ is the first measurement such that R_1 is the ratio prior to adding enzyme.

$$R_i = [\mathbf{77}]_i / [\mathbf{125}]_i = (h_{i,\mathbf{77}} \cdot \varepsilon_{278,\mathbf{125}}) / (h_{i,\mathbf{125}} \cdot \varepsilon_{278,\mathbf{77}})$$

$$f_i = 1 - \{(h_{i,\mathbf{77}} / \varepsilon_{278,\mathbf{77}}) + (h_{i,\mathbf{125}} \cdot \varepsilon_{278,\mathbf{125}})\} / \{(h_{1,\mathbf{77}} / \varepsilon_{278,\mathbf{77}}) + (h_{1,\mathbf{125}} \cdot \varepsilon_{278,\mathbf{125}})\} \cdot (h_{1,\mathbf{133}} / h_{i,\mathbf{133}})$$

The ratio of specificity constants (θ_1) can then be determined from the residual substrate ratio defined implicitly as a function of the fraction of reaction based on the model previously established in the study of kinetic isotope effects.[130] Fitting of this model to the experimental data was done by Dr. Mark W. Ruszczycky using implicit nonlinear least squares regression (*manuscript in preparation*).

$$\theta_1 = (V / K)_{\mathbf{77}} / (V / K)_{\mathbf{125}}$$
$$= \log[(1 - f) \cdot (1/R_1 + 1) / (1/R_f + 1)] / \log[(1 - f) \cdot (R_1 + 1) / (R_f + 1)]$$

3.2.7 Activity assay of ForS

The *N*-terminal His_6-tagged ForS was cloned and overexpressed as described in Chapter 3.2.5 (see Figure 3-5). The purification was conducted under the same conditions except the buffer composition. The lysis buffer was composed of 50 mM *N*-tris(hydroxymethyl)-methyl-2-aminoethanesulfonic acid (TES) (pH 7.5) containing 10 mM magnesium chloride. For Ni-NTA purification, the slurry was washed with the lysis buffer containing 20 mM imidazole and the protein was eluted with the lysis buffer containing 250 mM imidazole. The eluted protein was diluted prior to dialysis in order to prevent precipitation. The pooled protein was dialyzed against 1 L of 50 mM TES buffer (pH 7.5) containing 20% glycerol three times before storage at –80 °C.

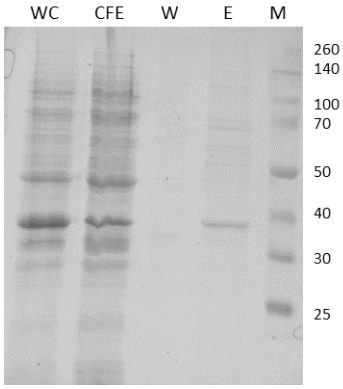


Figure 3-5. SDS-PAGE of ForS (expected molecular weight: 38.7 kDa). WC: whole cell, CFE: cell-free extract, W: wash, E: eluted protein, M: protein marker (Spectra multicolor broad range protein ladder).

To test the activity of ForS, a reaction mixture containing 1.25 mM PRPP (**74**), 5 mM of putative substrate in 50 mM TES buffer (pH 7.5) was incubated with 10 μM ForS at room temperature. The substrates tested were L-glutamate, D-glutamate, L-aspartate, α-ketoglutaric acid, (4*R*)-4-hydroxy-L-glutamate, 4-hydroxy-2-ketoglutaric acid, and 3,5-pyrazoledicarboxylic acid. All compounds are commercially available. The enzyme activity was assayed either by pyrophosphate assay or by LCMS. In order to quantitate pyrophosphate released upon coupling between PRPP (**74**) and substrate, the reaction mixture was quenched after 2 hours by boiling. After collecting supernatant by centrifugation, the concentration of pyrophosphate in 50 times diluted samples was determined according to the previously described procedure.[131] For LCMS analysis, the reaction mixture was deproteinized by using YM-10 centrifugal filtration and submitted for LCMS.

3.2.8 PCR-based Screening of *S. antibioticus*

The genomic DNA of *S. antibioticus* was extracted from the culture in TSB media grown at 30 °C for 2 days according to the reported procedure.[132] The extracted DNA was screened by PCR to identify the homologs of *hisG* and *hisA* using specific primers. Primers were designed based on known conserved sequences of *hisG* and *hisA* involved in the histidine biosynthesis in various *Streptomyces* species: 4 forward (hisG_1F to hisG_4F) and 3 reverse primers (hisG_1R to hisG_3R) for *hisG*, and 3 forward (hisA_1F to hisA_3F) and 2 reverse primers (hisA_1R and hisA_2R) for *hisA*.

hisG_1F: 5′-ATCGCCGTCCCCAACAAGGG-3′

hisG_2F: 5′-ATGCTSCATGAGGCSGGCTACC-3′

hisG_3F: 5′-AACGAGGTSGAGTTCTTCTACC-3′

hisG_4F: 5′-CTSGACATCGGCATCACSGG-3′

hisG_1R: 5′-GCAGTCGTAGTCCATCATCACG-3′

hisG_2R: 5′-ACCCAGCCCTCGTKGTGSAGC-3′

hisG_3R: 5′-TCGTCCATGACCCGCTGSGC-3′

hisA_1F: 5′-CGCCGTCGACGTCCGYGACG-3′

hisA_2F: 5′-CTSGTSCACGGCGWGTCCGG-3′

hisA_3F: 5′-GACCTSGACGCSGCSTTCGG-3′

hisA_1R: 5′-TGCCGTCCTTGGCGATGTCGG-3′

hisA_2R: 5′-TGAACGCCTTSGCGTASAGSGCC-3′

After PCR using the above primers and the genomic DNA as a template, the PCR fragments were cloned into the pCR-Blunt vector and sequenced. A total of 20 samples for *hisG* and 34 samples for *hisA* were sequenced. Only one *hisG* and one *hisA* were found.

3.2.9 Activity Assay of the Reductase Involved in the Biosynthesis of Coformycin

The gene *orf -2* encoded upsteam of the formycin A biosynthetic gene cluster was cloned and expressed as an *N*-terminal His_6-tagged protein and purified as described in Chapter 3.2.4 except for the buffer composition. The lysis buffer was 50 mM HEPES buffer (pH 8.0). The wash and elution buffers contained 25 mM or 250 mM imidazole, respectively. The purified protein was dialyzed against 50 mM HEPES (pH 8.0) containing 10% glycerol.

The activity of ORF -2 was tested by conducting a reverse reaction, which is the oxidation of 2′-deoxycoformycin (**109**) to 8-ketodeoxycoformycin (**120**). Specifically, 0.4 μM ORF -2 was incubated with 0.5 mM **109** in 50 mM HEPES buffer (pH 8.0) in the presence of 1 mM $NADP^+$ for 2 hours at room temperature. The UV absorbance at 340 nm was monitored to follow the production of NADPH. To quench the reaction, 4 volumes of ethanol was added and incubated at 4 °C for 10 min. The supernatant was collected by centrifugation and dried in a speed vacuum concentrator. The sample was re-dissolved in water and analyzed by HPLC using a Microsorb-MV 100-5 C18 column. The column was eluted with 2% acetonitrile in 50 mM ammonium acetate. 2′-Deoxycoformycin (**109**) was typically eluted at 19.3 min under these conditions. The peak at 16.1 min was isolated and analyzed by ESI-MS in positive ion mode. Its molecular weight is consistent with that of the oxidized product 8-ketodeoxycoformycin (**120**); calculated for $C_{11}H_{15}O_4N_4Na^+$ [M + $Na]^+$: 289.0913 m/z, observed: 289.1 m/z.

3.3 Results and Discussion

3.3.1 Identification of the Putative Biosynthetic Gene Cluster of Formycin A

As the genomic information of a formycin-producing strain is not available, the genome of *S. kaniharaensis* was sequenced in order to identify the biosynthetic pathway of formycin A. Over 3000 contiguous sequences (contigs) with an average size of 2641 bp were obtained from genome sequencing. Although the size of the contigs was relatively short for thorough inspection of the biosynthetic gene cluster, the preliminary analysis revealed that two homologous copies of the *purA*, *purB*, *purC*, and *purH* genes reside in *S. kaniharaensis*. The enzymes encoded by those *pur* genes (PurA, PurB, PurC, and PurH) are reported to catalyze the conversion of carboxyaminoimidazole ribonucleotide (**121**, CAIR) to AMP (**127**) during the later stage of adenosine biosynthesis (Figure 3-6).[133,134] The second set of *pur* homologous genes exhibited relatively low similarity compared to the highly conserved *pur* genes involved in adenosine biosynthesis, which allowed the distinction between these two sets of genes (see Table 3-3). Given the considerable structural resemblance of formycin A (**51**) and adenosine (**53**), the presence of a second set of *pur* genes was indicative of an analogous biosynthetic pathway underlying assembly of the pyrazolopyrimidine nucleobase in formycin A.

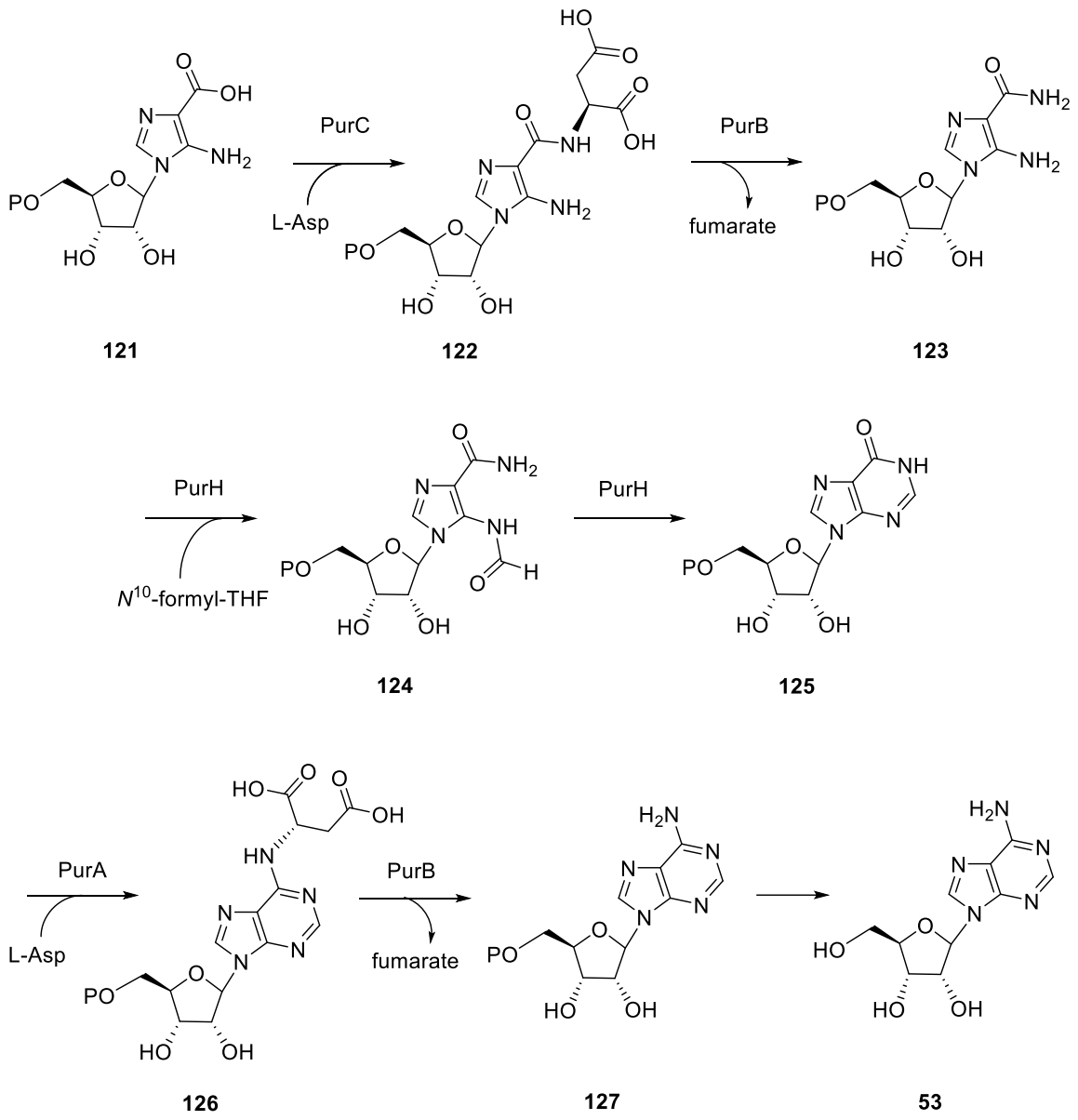


Figure 3-6. The later steps in the biosynthesis of adenosine (**53**).

Enzymes involved in the production of bacterial secondary metabolites are commonly encoded by chromosomally clustered genes (*i.e.*, biosynthetic gene clusters).[135] In an effort to discover the entire gene cluster for formycin A, a cosmid library was constructed from the genomic DNA of *S. kaniharaensis* and screened by PCR. The primers used for PCR were specifically designed based on the sequences of the second copy of *purA*, *purC*, and *purH* genes. A single cosmid, designated K24C, was found to harbor all of the *purA*, *purC*, and *purH* homologous genes, while the K5B cosmid was found to contain the *purC* homologous gene. The cosmid K24C was fully sequenced by primer walking, and it was found that the size of the K24C cosmid is around 41 kbp containing 35 ORFs defined within the cosmid. The functional annotation based on bioinformatics analysis using a BLAST search permitted identification of the biosynthetic gene cluster of formycin A in the K24C cosmid (Figure 3-7). Spanning approximately 30 kbp, the putative formycin A gene cluster contains 24 ORFs. The complete functional annotation of the individual gene products in the putative gene cluster are shown in Table 3-2. The second copy of homologous *purA* (*forD*), *purB* (*forR*), *purC* (*forV*), and *purH* (*forC*) found in the gene cluster exhibits 50% sequence similarities with the *pur* genes involved in the biosynthesis of adenosine in *S. kaniharaensis* (Table 3-3). On the contrary, the genuine *pur* genes for the essential primary pathway are highly conserved in all *Streptomyces* species showing 90% sequence similarities. The differences between these sets of genes strongly suggest that the second set of genes is involved in the secondary metabolism and may have evolved from the primary biosynthetic pathway. It should be also noted that significant similarities were found between the genes in the identified *for* gene cluster and those encoded in uncharacterized gene clusters in *Streptomyces resistomycificus*[136] and *Salinispora arenicola*,[137] although whether the latter two strains produce formycin A or a derivative thereof is not known.

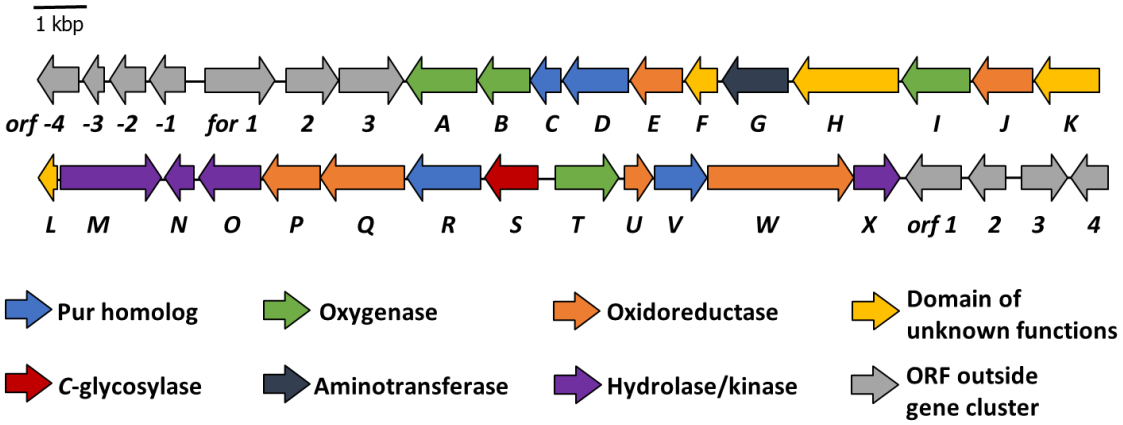


Figure 3-7. Biosynthetic gene cluster of formycin A in *S. kaniharaensis*. The putative functions of each gene product are color-coded, and the ORFs considered outside the gene cluster are shown in gray. The complete functional annotation is listed in Table 3-2.

Table 3-2. Functional annotation of individual gene product in formycin A biosynthetic gene cluster.

Gene	Size (aa)	Protein Homolog [Origin]	Accession	Identity/ Similarity (%/%)
for1	453	GntR family transcriptional regulator [*Streptomyces resistomycificus*]	WP_037802186.1	72/79
for2	336	Hypothetical protein [*Streptomyces resistomycificus*]	WP_030038112.1	34/47
for3	415	Major facilitator superfamily (MFS) transporter [*Streptomyces resistomycificus*] MFS transporter [*Salinispora arenicola*]	WP_053190963.1 WP_012180823.1	65/78 49/61
forA	450	Xenobiotic compound monooxygenase A subunit [*Salinispora arenicola*]	WP_012180830.1	79/88
forB	343	Hypothetical protein [*Streptomyces resistomycificus*] Putative monooxygenase [*Salinispora arenicola*]	WP_051869545.1 ABV96521.1	60/70 57/67
forC	197	Phosphoribosylaminoimidazolecarboxamide formyltransferase [*Streptomyces resistomycificus*] Phosphoribosylaminoimidazolecarboxamide formyltransferase [*Salinispora arenicola*]	WP_030038113.1 WP_028673470.1	76/85 65/72

Table 3-2. Continued.

forD	424	Adenylosuccinate synthetase [*Streptomyces resistomycificus*] Adenylosuccinate synthetase [*Salinispora arenicola*]	WP_030038114.1 WP_029537068.1	73/84 64/77
forE	334	2-hydroxyacid dehydrogenase, partial [*Streptomyces resistomycificus*] D-Isomer specific 2-hydroxyacid dehydrogenase [*Salinispora arenicola*]	WP_030038115.1 WP_012180805.1	84/91 70/81
forF	211	Hypothetical protein [*Streptomyces resistomycificus*]	WP_051869535.1	69/75
forG	424	Membrane protein [*Streptomyces resistomycificus*] Class III aminotransferase [*Salinispora arenicola*]	WP_030038117.1 WP_012180807.1	83/91 68/77
forH	675	Methionine–tRNA ligase, partial [*Streptomyces resistomycificus*] Methionine–tRNA ligase [*Salinispora arenicola*]	WP_051869537.1 WP_029537071.1	67/75 61/73
forI	441	L-Lysine 6-monooxygenase [*Streptomyces resistomycificus*] L-Lysine 6-monooxygenase [*Salinispora arenicola*]	WP_030038120.1 WP_028675079.1	72/84 63/78
forJ	385	Saccharopine dehydrogenase [*Streptomyces resistomycificus*] Saccharopine dehydrogenase [*Salinispora arenicola*]	WP_053190969.1 WP_028676437.1	65/76 58/70

Table 3-2. Continued.

forK	421	Phosphoribosylglycinamide synthetase [*Streptomyces resistomycificus*] Phosphoribosylglycinamide synthetase [*Salinispora arenicola*]	WP_030038122.1 WP_012180811.1	81/88 69/80
forL	128	Hypothetical protein		
forM	646	Hypothetical protein [*Streptomyces resistomycificus*] Peptidase S9 [*Salinispora arenicola*]	WP_053190961.1 WP_028678644.1	58/68 53/65
forN	191	NUDIX hydrolase [Streptomyces] NUDIX hydrolase [*Salinispora arenicola*]	WP_030290725.1 WP_016815025.1	82/87 60/71
forO	397	Amidohydrolase [*Streptomyces resistomycificus*] Amidohydrolase [*Salinispora arenicola*]	WP_053190965.1 WP_012180815.1	83/89 70/79
forP	375	Hypothetical protein [*Streptomyces resistomycificus*] FAD-dependent oxidoreductase [*Salinispora arenicola*]	WP_030038132.1 WP_012180814.1	73/79 59/67
forQ	539	Fumarate reductase [*Streptomyces resistomycificus*] Fumarate reductase [*Salinispora arenicola*]	WP_030038133.1 WP_038842037.1	77/84 62/71
forR	475	Adenylosuccinate lyase [*Streptomyces resistomycificus*]	WP_030038134.1	82/89

Table 3-2. Continued.

forS	342	*β*-Ribofuranosylaminobenzene 5′-phosphate synthase [*Streptomyces resistomycificus*] *β*-Ribofuranosylaminobenzene 5′-phosphate synthase [*Salinispora arenicola*]	WP_030038135.1 WP_012180817.1	78/84 67/75
forT	409	Phthalate 4,5-dioxygenase [*Streptomyces resistomycificus*] Phthalate 4,5-dioxygenase [*Salinispora arenicola*]	WP_053190967.1 WP_028677418.1	87/92 75/85
forU	186	FMN reductase [Streptomyces]	WP_030247495.1	75/85
forV	334	SAICAR synthetase [*Streptomyces resistomycificus*] SAICAR synthetase [*Salinispora arenicola*]	WP_051869569.1 WP_028673475.1	65/75 59/70
forW	937	Oxidoreductase [Mycobacterium]	WP_024454584.1	62/72
forX	296	Ribokinase [*Salinispora arenicola*] Ribokinase [*Streptomyces resistomycificus*]	WP_029537066.1 WP_051869540.1	57/66 55/67

Table 3-3. Sequence similarity between two sets of Pur-like enzymes.

	Identity/Similarity (%/%)			
S. kaniharaensis	*S. coelicolor*	*S. griseus*	*S. avermitilis*	*S. scabiei*
	PurA (427 aa)	PurA (427 aa)	PurA (427 aa)	PurA (427 aa)
PurA (427 aa)	88/94	88/95	86/94	87/93
ForD (423 aa)	40/56	39/56	40/57	40/57
	PurC (299 aa)	PurC (304 aa)	PurC (299 aa)	PurC (299 aa)
PurC (299 aa)	75/87	75/84	76/87	76/87
ForV (333 aa)	31/41	28/43	30/40	29/42
	PurH (523 aa)	PurH (517 aa)	PurH (527 aa)	PurH (520 aa)
PurH (502 aa)	83/89	82/87	81/88	82/89
ForC (196 aa)	40/53	38/51	38/53	39/53
	PurB (480 aa)	PurB (480 aa)	PurB (479 aa)	PurB (480 aa)
PurB* (336 aa)	85/92	85/93	85/92	85/92
ForR (474 aa)	74/85	74/85	74/85	75/85

* Only partial sequence (336 aa) is known.

ORFs flanked by the putative gene cluster at both ends (For1, 2, 3 and ORFs 1, 2, 3, 4 in Figure 3-7) encode a transcriptional regulator, a transporter, enzymes involved in primary metabolism, or proteins of unknown function, and are likely the boundaries of the cluster. *E. coli* transformed with the K24C cosmid showed reduced lifespans, and the cosmid isolated after passaging the cells was found to lose part of the putative gene cluster. Heterologous expression of the biosynthetic gene cluster was hindered due to the partial loss of the gene cluster. However, those observations may suggest that the metabolite produced by the enzymes encoded in the putative gene cluster are toxic to the host bacteria *E. coli*.

3.3.2 Formation of Pyrazolopyrimidine Nucleobase

As discussed in the previous section, the presence of *pur*-like genes including *forD*, *forR*, *forV*, and *forC* implied that the later steps of formycin A biosynthesis are analogous to those of the adenosine biosynthetic pathway (Figure 3-8). Each enzyme was overexpressed in *E. coli*, and three pathway intermediates were synthesized. The enzyme activities were assayed to test the proposed pathway for the assembly of pyrimidine ring of formycin A.

ForV and ForR are proposed to catalyze the conversion of carboxylic acid **128** to amide **130**. The consumption of carboxyaminopyrazole ribonucleotide (**128**) was observed by HPLC when 0.5 mM **128** was incubated with 20 μM ForV in the presence of 10 mM L-aspartate and 2 mM ATP (Figure 3-9, traces 1 and 2). While no peak corresponding to **129** was detected due to the prolonged retention of **129** in the column under this condition, the formation of ADP indicated reaction progress. The addition of 20 μM ForR to the reaction mixture resulted in a new peak at the retention time of 7.8 min (Figure 3-9, trace 3), which was identified to be 5-aminopyrazole-4-carboxamide ribonucleotide (**130**) based on ESI-MS analysis (calculated for $C_9H_{14}O_8N_4P^-$ $[M-H]^-$: 337.0555 *m*/*z*, found: 337.1 *m*/*z*). The product also co-eluted with a chemically synthesized standard **130**. These results showed that the amide moiety of **130** is introduced by ForV and ForR using L-aspartate as the amino donor.

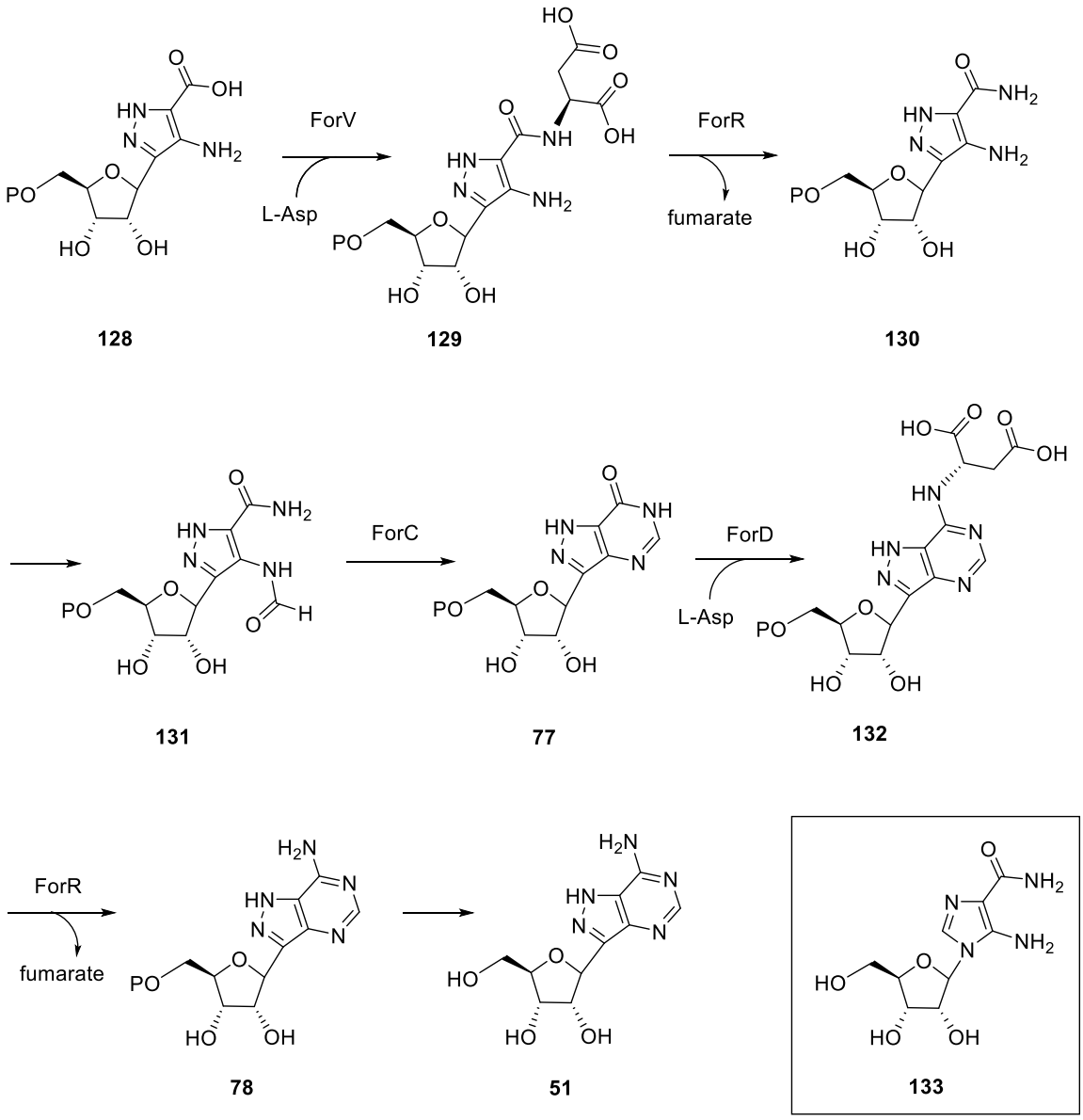


Figure 3-8. The later steps of proposed pathway for formycin A (**51**) are analogous to those in the primary adenosine biosynthetic pathway. The structure of 5-aminoimidazole-4-carboxamide 1-*β*-D-ribofuranoside (**133**, AICAR) used as a HPLC internal standard in competition experiment (see Chapter 3.3.3) is shown in the box.

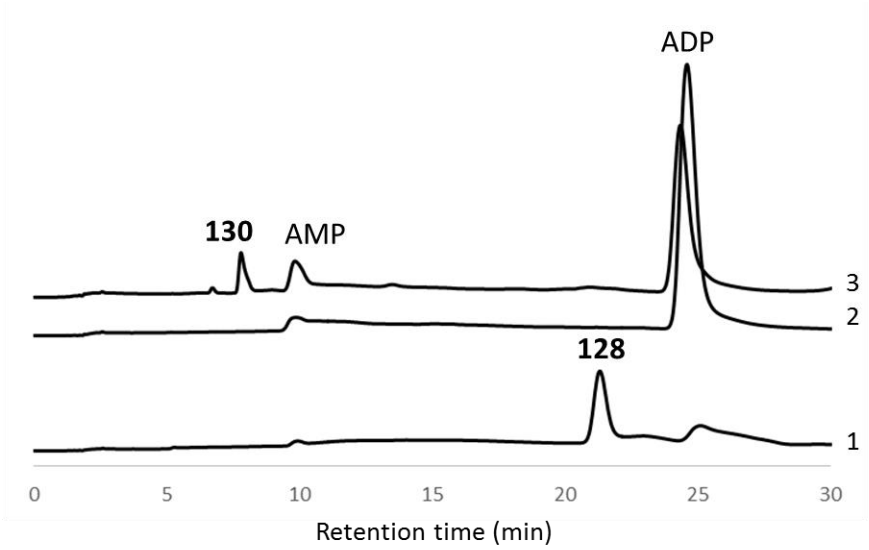


Figure 3-9. HPLC traces showing the reactions catalyzed by ForV and ForR. Trace 1 is the reaction mixture of **128**, L-aspartate, and ATP analyzed before addition of enzymes. Trace 2 is the same reaction mixture as trace 1 incubated with ForV for 4 h. Trace 3 is the same reaction mixture as trace 1 incubated with ForV and ForR together for 4 h. It was noted that AMP was formed in the reaction mixture for unknown reason.

The bacterial PurH involved in the biosynthesis of adenosine is a bifunctional enzyme that consists of the *N*-terminal IMP cyclohydrolase domain and the *C*-terminal AICAR formyltransferase domain.[129,134] It should be pointed out that ForC in the formycin A gene cluster is less than half of PurH in size, and has homology only to the *N*-terminal domain of PurH. As the formyltransferase domain is missing in ForC, it did not react with 5-aminopyrazole-4-carboxamide ribonucleotide (**130**). Instead, PurH of *S. kaniharaensis* exhibited the formyltransferase activity when 30 μM PurH was anaerobically incubated with 0.5 mM **130** and 1 mM N^{10}-formyl-THF (Figure 3-10, traces 1 and 2). The isolated product eluting at 7.4 min was found to be 5-formaminopyrazole-4-carboxamide ribonucleotide (**131**), which gave signals at 365.1 and 387.0 *m*/*z* corresponding to $[M-H]^-$

and $[M - 2H + Na]^-$ ions of **131** in ESI-MS. The reaction of *E. coli* PurH with **130** also produced the formylated compound **131**, rather than the cyclized product **77**. This was in contrast with the formation of inosine monophosphate (**125**, IMP) from the reaction of *E. coli* PurH with 5-aminoimidazole-4-carboxamide ribonucleotide (**123**) (see Figure 3-6). It is noteworthy that *E. coli* PurH does not catalyze the cyclization to form formycin B 5′-phosphate (**77**), which alludes to the specificity of Pur enzymes towards two close structural isomers. It is not clear as of now whether the formyl transfer reaction is catalyzed by an enzyme encoded in the putative gene cluster. When both 50 μM ForC and 30 μM PurH were included in the anaerobic incubation, ForC was capable of catalyzing the cyclization of the formylated intermediate **131**. The product, formycin B 5′-phosphate (**77**), was detected as $[M - H]^-$ and $[M - 2H + Na]^-$ at 347.0 and 369.0 *m*/*z*, respectively.

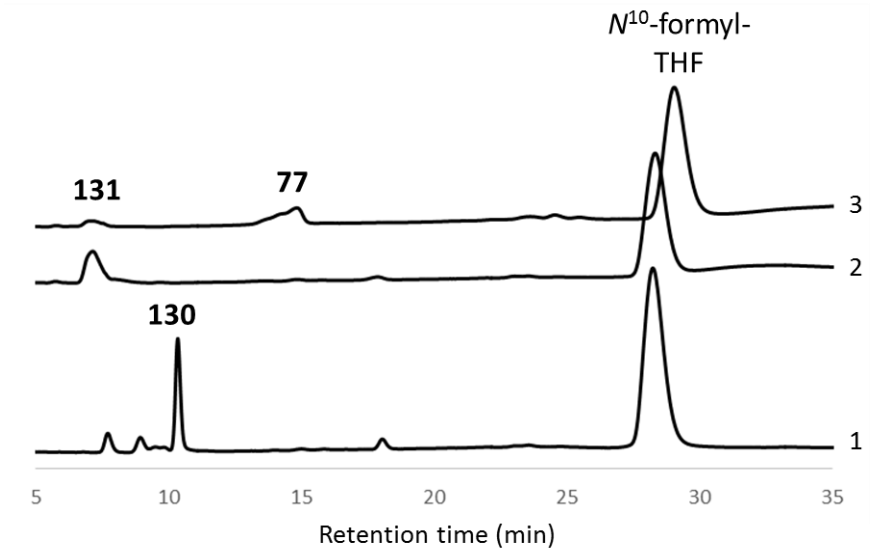


Figure 3-10. HPLC traces showing the reaction of PurH and ForC. Trace 1 is the reaction mixture containing **130** and N^{10}-formyl-THF before addition of enzymes. Trace 2 is the same reaction mixture as trace 1 incubated with PurH of *S. kaniharaensis* for 1 h. Trace 3 is the same reaction mixture as trace 1 incubated with PurH of *S. kaniharaensis* and ForC for 1 h. All reactions were conducted anaerobically inside glove box.

Formycin A 5′-phosphate (**78**) is proposed to be biosynthesized from formycin B 5′-phosphate (**77**) by ForD and ForR, which is consistent with an early experiment showing **77** as a likely precursor to formycin A (see Figure 3-8).[66] The consumption of formycin B 5′-phosphate (**77**) was observed upon the addition of 10 μM ForD into the reaction mixture of 1 mM **77**, 2.5 mM L-aspartate, and 0.125 mM GTP (Figure 3-11, traces 1 and 2). The formation of succinylformycin 5′-phosphate (**132**) was demonstrated by high resolution LC-MS analysis (calculated for $C_{14}H_{17}N_5O_{11}P^-$ $[M-H]^-$: 462.0668 *m*/*z*, found: 462.0658 *m*/*z*), which was also observable by HPLC using a Microsorb-MV 100-5 C18 column (see Figure 3-12). The coupling reaction of 10 μM ForD and 10 μM ForR produced a new peak with the retention time of 14.6 min (Figure 3-11, trace 3). The new product was identified to be formycin A 5′-phosphate (**78**) by ESI-MS analysis, which gave a signal at 346.0556 *m*/*z* corresponding to $[M-H]^-$ ion of **78** (calculated for $C_{10}H_{13}N_5O_7P^-$ $[M-H]^-$: 346.0558 *m*/*z*).

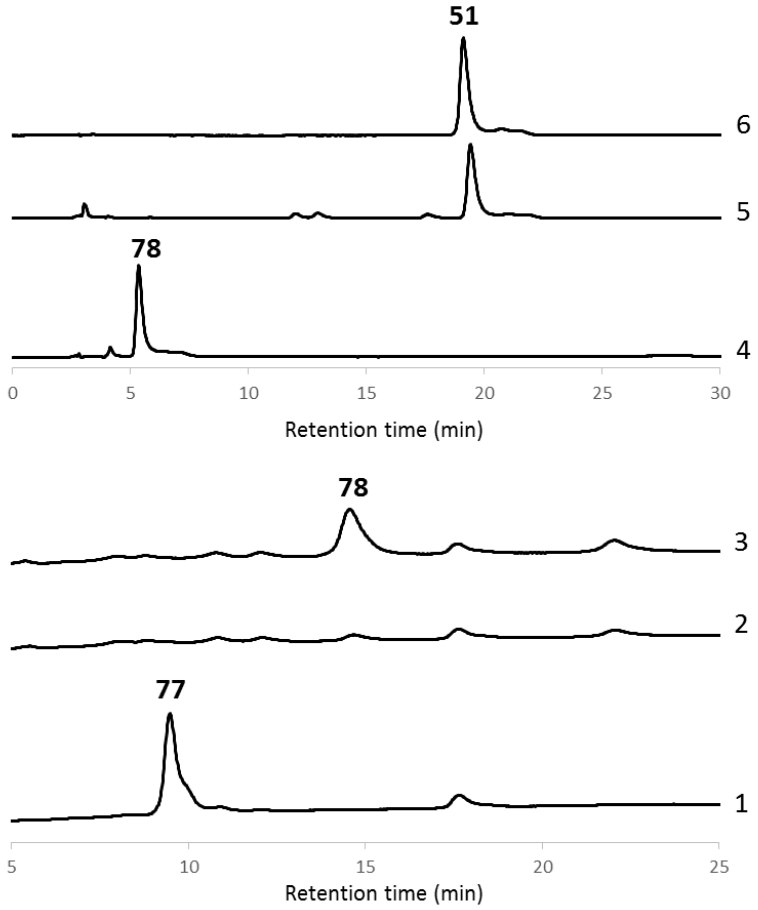


Figure 3-11. HPLC traces showing the reactions catalyzed by ForD and ForR. Trace 1 is the reaction mixture containing **77**, L-aspartate, and GTP, which was analyzed before addition of enzymes. Trace 2 is the same reaction mixture as trace 1 incubated with ForD for 2 h. Trace 3 is the same reaction mixture as trace 1 incubated with ForD and ForR together for 2 h. Trace 4 is the reaction product **78** isolated from trace 3. Trace 5 is the incubation of **78** with CIP for 1 h that resulted in the formation of **51**. Trace 6 is the injection of standard formycin A, which shows that the retention time of **51** is the same as that of formycin A. Traces 1 to 3 were analyzed by HPLC anion exchange column, whereas traces 4 to 6 were analyzed by C18 column (see Chapter 3.2.5.4 for detailed HPLC methods).

The identity of the ForD/ForR reaction product as formycin A 5′-phosphate (**78**) was further confirmed by hydrolyzing the phosphate moiety. Incubation of the isolated reaction product with 100 units of CIP resulted in a new compound that co-eluted with the commercial standard of formycin A (**51**) (see Figure 3-11, traces 5 and 6). The mass signal at 266.0893 *m*/*z* observed by high resolution ESI-MS was consistent with $[M - H]^-$ ion of formycin A (calculated for $C_{10}H_{12}N_5O_4^-$ $[M - H]^-$: 266.0895 *m*/*z*). These results demonstrated the conversion of formycin B 5′-phosphate (**77**) to formycin A 5′-phosphate (**78**) by ForD and ForR as proposed. Taken together, the pyrimidine ring in formycin A is assembled by catalysis of Pur-like enzymes including ForV, ForR, ForC, and ForD in an analogous pathway to adenosine biosynthesis, which supports the hypothesis that the biosynthesis of formycin A is carried out by enzymes encoded in the identified biosynthetic gene cluster (see Figure 3-8).

3.3.3 Specificity of ForD

The reactions catalyzed by Pur-like enzymes during construction of the pyrozolopyrimidine nucleobase in formycin A are likely to follow similar reaction mechanisms compared to the corresponding reactions in the adenosine biosynthetic pathway. This prompted the question as to whether the homologous enzymes exhibit specificity/promiscuity toward putative intermediates in formycin A and adenosine pathways.

In order to examine the substrate specificity, the reaction of ForD and PurA with either formycin B 5′-phosphate (**77**) or IMP (**125**) as a substrate were investigated. It was shown that ForD can accept both **77** and **125** as a substrate (Figure 3-12). The condensation

reaction between **77** and L-aspartate by ForD resulted in the formation of succinylformycin 5′-phosphate (**132**) (traces 1 to 3). Likewise, the production of succinyladenosine 5′-monophosphate (**126**, SAMP) was observed from the ForD reaction of **125** and L-aspartate (traces 4 to 6). On the other hand, PurA was shown to strictly recognize **125** and did not react with **77** (Figure 3-13).

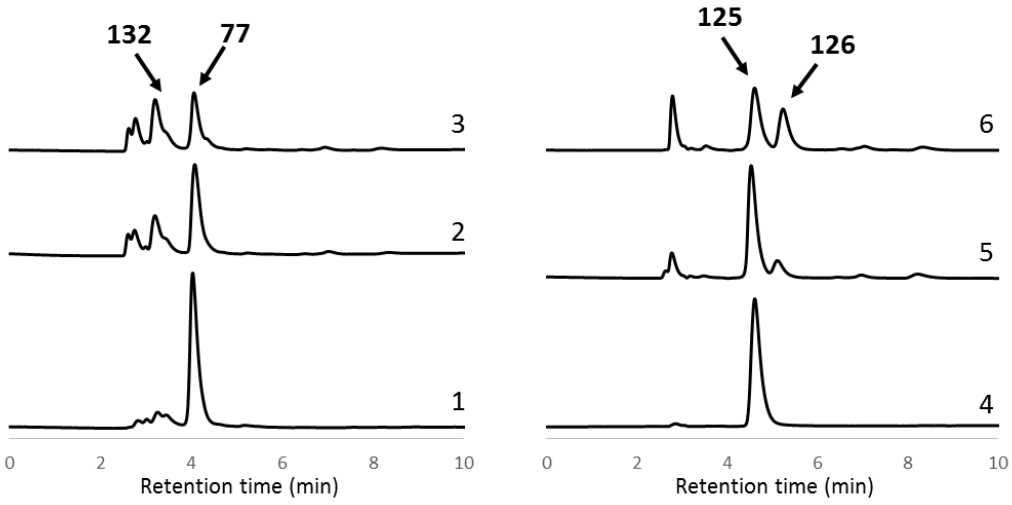


Figure 3-12. HPLC traces of ForD reactions. Traces 1 to 3 are the ForD reaction mixture with **77** as a substrate. Trace 1 was measured before adding enzymes, and traces 2 and 3 were measured after 10 min and 25 min upon addition of ForD. Traces 4 to 6 are the ForD reaction mixture with **125**. Trace 4 was measured before adding enzymes, and traces 5 and 6 were measured after 5 min and 25 min upon addition of ForD. The peak at 2.8 min is an unknown contaminant from enzyme preparation.

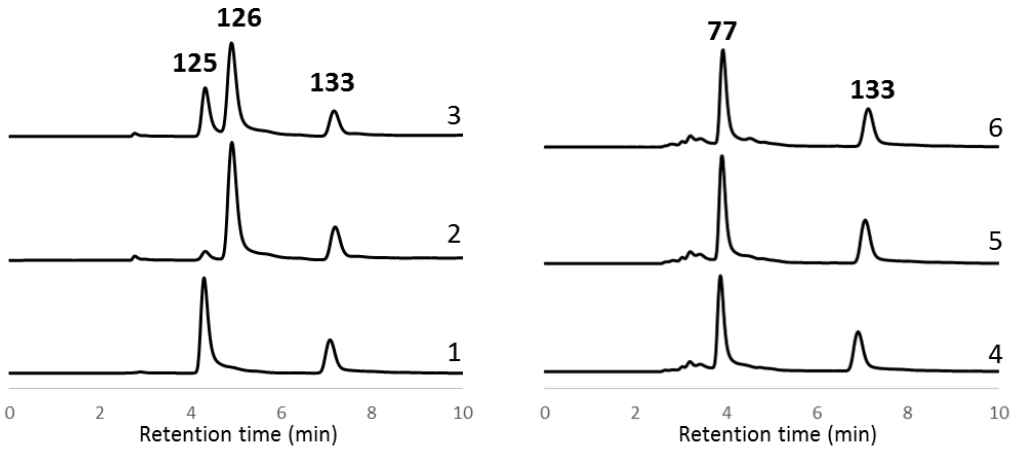


Figure 3-13. HPLC traces of PurA reactions. Trace 1 is the reaction mixture with **125** before adding enzyme, and trace 2 is the same reaction mixture incubated with PurA for 10 min. Trace 3 is the co-injection of trace 2 and **125**. Traces 4 to 6 are the PurA reaction mixture with **77**. Trace 4 was measured before adding enzymes, and traces 5 and 6 were measured after 10 min and 60 min upon addition of ForD. AICAR (**133**) was included as an internal standard.

While ForD was able to process both substrates, it appeared that ForD indeed prefers formycin B 5′-phosphate (**77**) over IMP (**125**), which is consistent with **77** being the biosynthetically relevant substrate for ForD (Figure 3-14). A more quantitative analysis was provided by a competition experiment in which the reaction progress was monitored in the presence of both substrates. The samples were taken prior to addition of ForD and at given times, and the concentration ratio of **77** to **125** (R_i) was determined based on the HPLC peak integration and extinction coefficients. The fraction of reaction (f_i) was calculated by taking relative integration of each peak to the area of internal standard AICAR (**133**, see Figure 3-8 for the structure). The substrate enrichment was observed as one of the substrates (*i.e.*, formycin B 5′-phosphate (**77**)) was consumed faster over the

other (Figure 3-15). The specificity constant ratio (θ_1) was then found to be 8.65 ± 0.58 by averaging estimates obtained from implicit nonlinear least squares regression over five separate trials. These results indicate that ForD preferentially recognizes and processes **77** over **125** despite their very similar structures, which are in a good agreement with the functional assignment of Pur-like enzymes to the biosynthesis of formycin A.

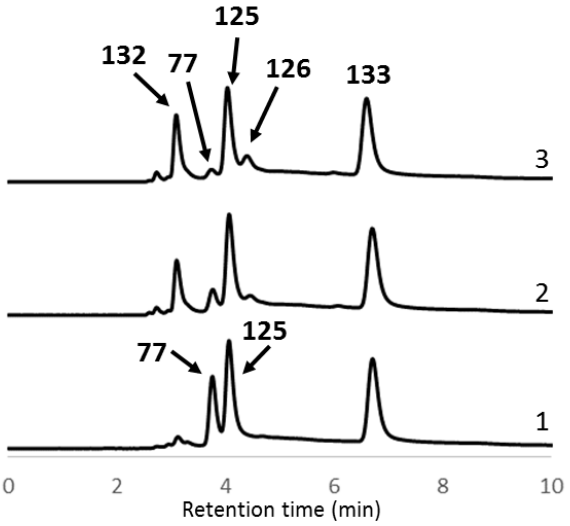


Figure 3-14. Competition experiment. Trace 1 is the reaction mixture with **77** and **125** measured before adding ForD. Traces 2 and 3 were measured after 10 min and 20 min upon addition of ForD. AICAR (**133**) was included as an internal standard.

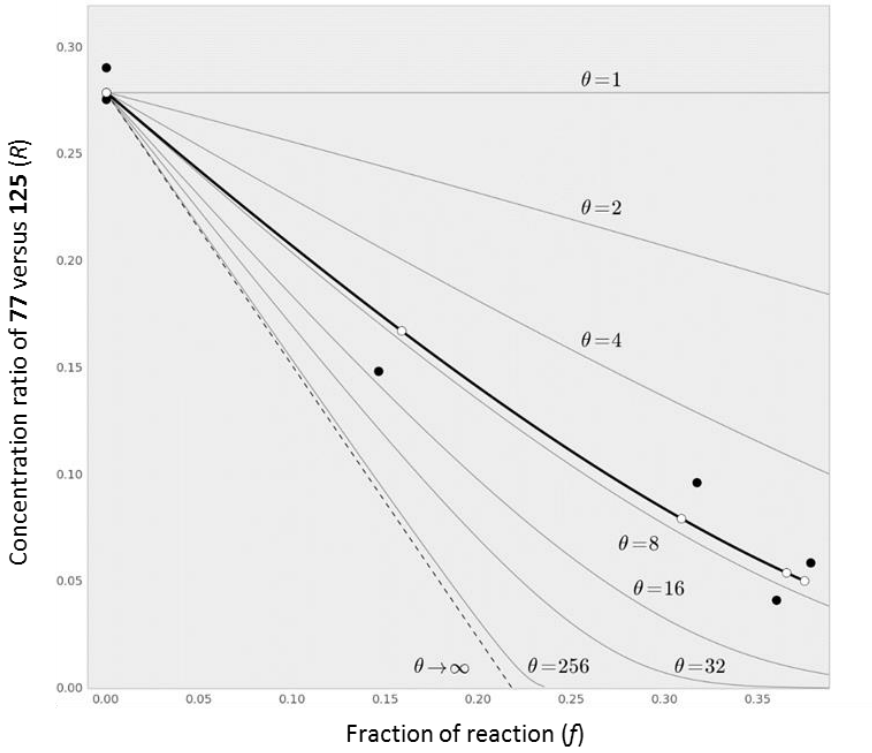


Figure 3-15. The representative plot of substrate concentration ratio of **77** to **125** (R_i) as a function of fraction of reaction (f_i). The solid black circles are the observed experimental data. The white circles and black line denote the corresponding fitted data points. The gray lines are the calculated curves of R versus f at different fixed specificity constant ratios (θ_1). The broken line denotes the asymptote for $\theta_1 \rightarrow \infty$.

3.3.4 Proposed Biosynthesis of Formycin A

As discussed in Chapter 3.3.2, the later steps of formycin A biosynthesis are expected to follow a pathway analogous to that of adenosine biogenesis. The biotransformation from carboxyaminopyrazole ribonucleotide (**128**) to formycin A 5′-phosphate (**78**) in formycin A biosynthesis was shown to be catalyzed by a series of Pur-like enzymes—ForV, ForR, ForC, and ForD (see Figure 3-8). However, the early pathway leading to the formation of **128** seems to be distinct from the biosynthesis of CAIR (**121**) and remains speculative at the present time. Nevertheless, the construction of a *C*-glycosidic bond between PRPP (**74**) and a glutamate derivative is proposed to take place at the early stage considering the asymmetrical incorporation pattern of L-glutamate (**75**) as demonstrated by feeding experiments, because the isotope labeling in the final product would be scrambled otherwise (see Chapter 1.2.3).[69,70] L-Glutamate (**75**) is expected to be incorporated as a single unit, which also suggests that the N7 nitrogen of formycin A is likely derived from the *α*-amino group of L-glutamate (**75**). The original proposal in which the N7 nitrogen originates from the *ε*-amino group of L-lysine (**76**) (see Figure 1-20) is likely due to *in vivo* catabolism of **76** to **75**.[138]

Interestingly, the gene product of *forS* exhibits 31% sequence identity and 47% similarity to MJ1427, a *β*-ribofuranosylaminobenzene 5′-phosphate synthase found in *Methanocaldococcus jannaschii*. MJ1427 catalyzes the decarboxylative coupling of PRPP (**74**) with 4-hydroxybenzoic acid (**134**) during the first step of methanopterin biosynthesis (Figure 3-16).[139,140] The considerable homology of ForS to MJ1427 implied that a similar coupling reaction of PRPP (**74**) with L-glutamate (**75**) or a derivative thereof catalyzed by ForS could result in the *C*-glycosidic linkage.

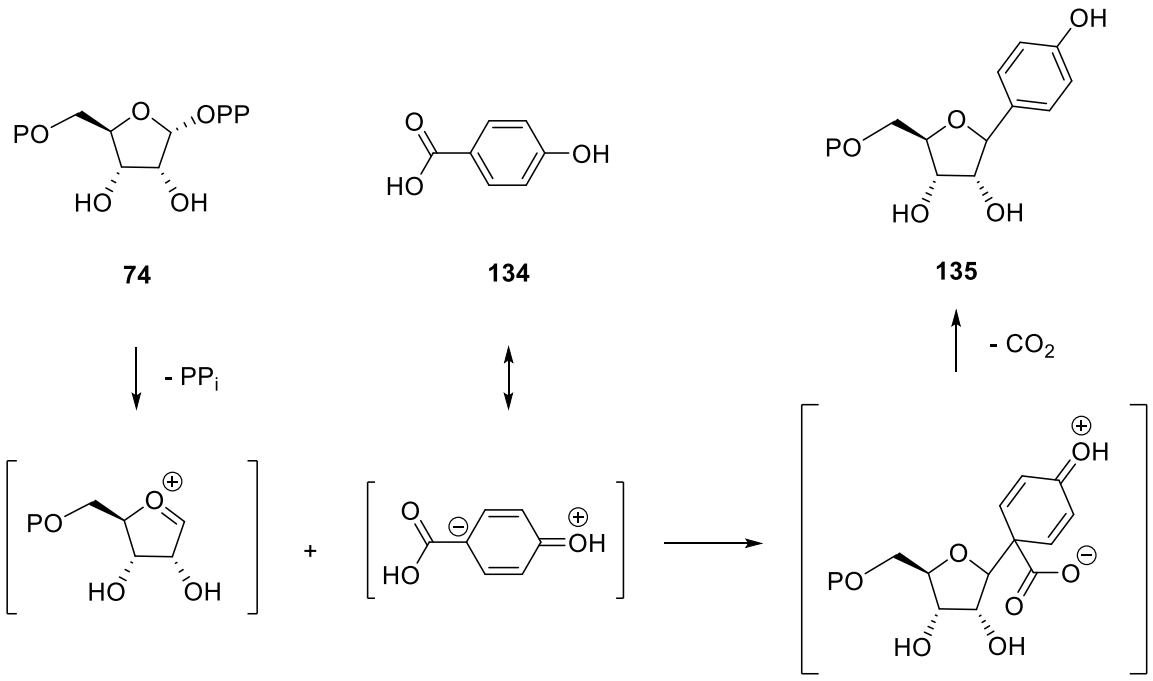


Figure 3-16. Formation of *C*-glycosidic bond by MJ1427 in *M. jannaschii*.

The *forS* gene was heterologously overexpressed as an *N*-terminal His_6-tagged protein in *E. coli*. The activity of ForS was assayed either by quantitating pyrophosphate released upon the formation of the *C*-glycosidic bond or by LCMS analysis. The substrates tested using the pyrophosphate assay included L-glutamate, D-glutamate, and *α*-ketoglutaric acid; however, none of them showed any activity (Figure 3-17). Likewise, the analysis by LCMS showed that the incubation of ForS with (4*R*)-4-hydroxy-L-glutamate (**136**) or 4-hydroxy-2-ketoglutaric acid (**137**) did not result in a product (Figure 3-18). These observations implied that the substrate for ForS has to be activated prior to nucleophilic addition to PRPP (**74**). L-Glutamate (**75**) itself does not have any functionality to stabilize the anionic character required for nucleophilic attack. On the contrary, the substrate of MJ1427 (**134**) has the aromatic ring with a *para* substituent that promotes the

decarboxylative coupling in a similar mechanism to the aromatic substitution. It should be mentioned that there was no reaction when 3,5-pyrazoledicarboxylic acid (**138**) was incubated with ForS, which suggested formation of the *C*-glycosidic linkage prior to assembly of the pyrazolopyrimidine moiety.

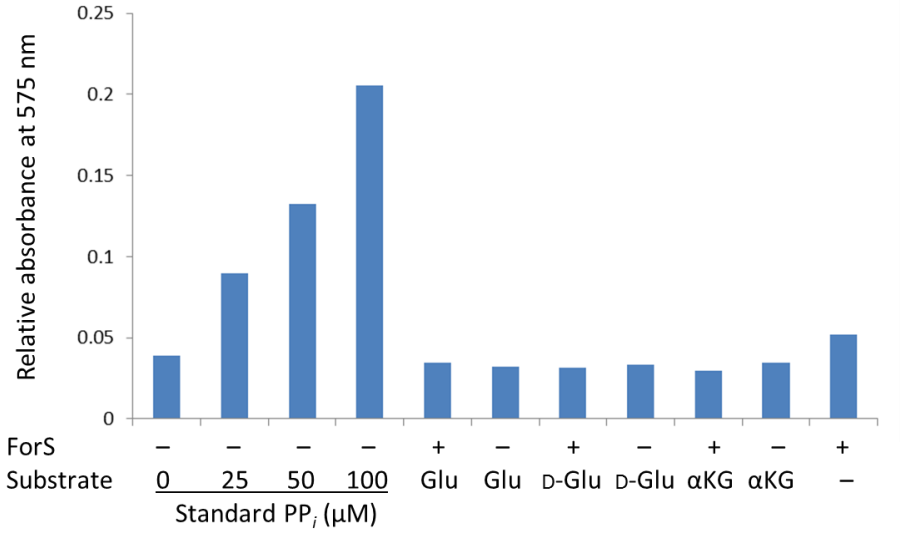


Figure 3-17. Activity assay for the formation of *C*-glycosidic bond by ForS.

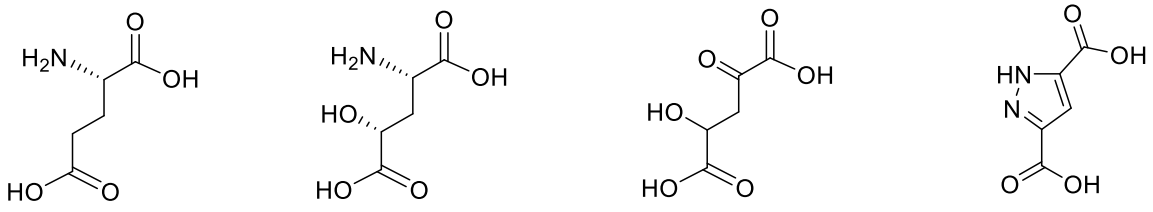


Figure 3-18. Structures of L-glutamate (**75**) and its derivatives tested in ForS assay.

The early part of the biosynthetic pathway for formycin A is proposed based on functional annotation of the enzymes found in the gene cluster, as shown in Figure 3-19 and Figure 3-20. Although the substrate for ForS has not been identified, the *C*-glycosydic linkage in formycin A seems to be constructed in a reaction catalyzed by ForS that facilitates nucleophilic addition of a derivative of L-glutamate to PRPP (**74**). In this respect, *β*,*γ*-unsaturated glutamate (**139**) or *β*-keto glutamate (**152**) is proposed as the potential substrate for ForS. ForQ is a homolog of succinate dehydrogenase, which may catalyze the dehydrogenation of **75** preceding the C–C bond formation (see Figure 3-19). The formation of a *C*-glycosydic linkage between **139** and **74** is then catalyzed by ForS, facilitated by the anionic character developed upon *α*-deprotonation and delocalized over the conjugating system. Subsequent dihydroxylation and oxidation by ForT and ForE, respectively, can produce the intermediate **142**. The alternative pathway proposes *β*-hydroxylation of **75** by ForA or ForB to be the first step, followed by the oxidation of hydroxyl group by ForE to form *β*-keto glutamate (**152**) (Figure 3-20). The decarboxylative coupling by ForS may generate **153**. Then hydroxylation of **153** by ForA or ForB may lead to the intermediate **154**, which is readily tautomerized to **142**.

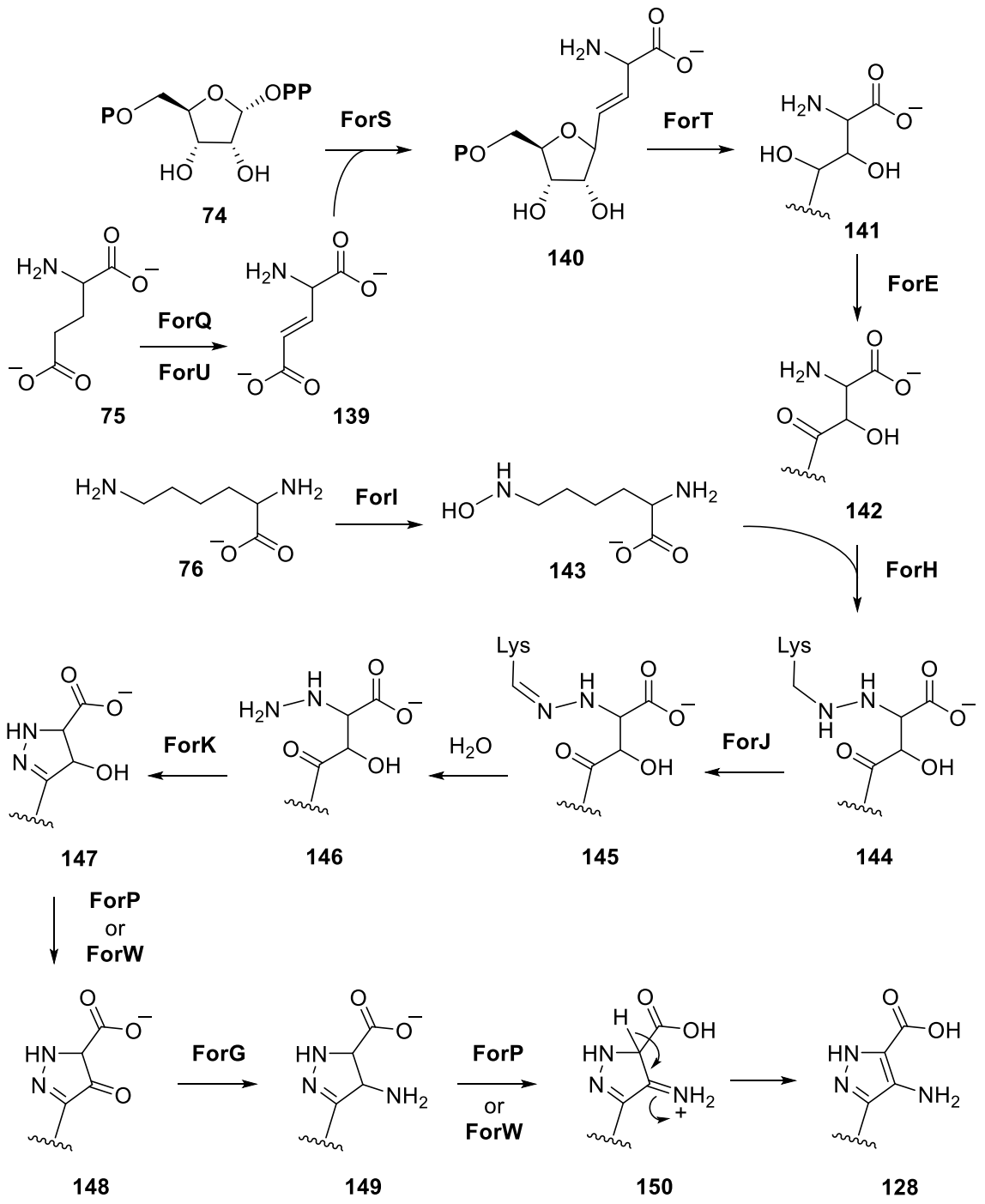


Figure 3-19. Proposed biosynthetic pathway for the formation of putative intermediate, carboxyaminopyrazole ribonucleotide (**128**).

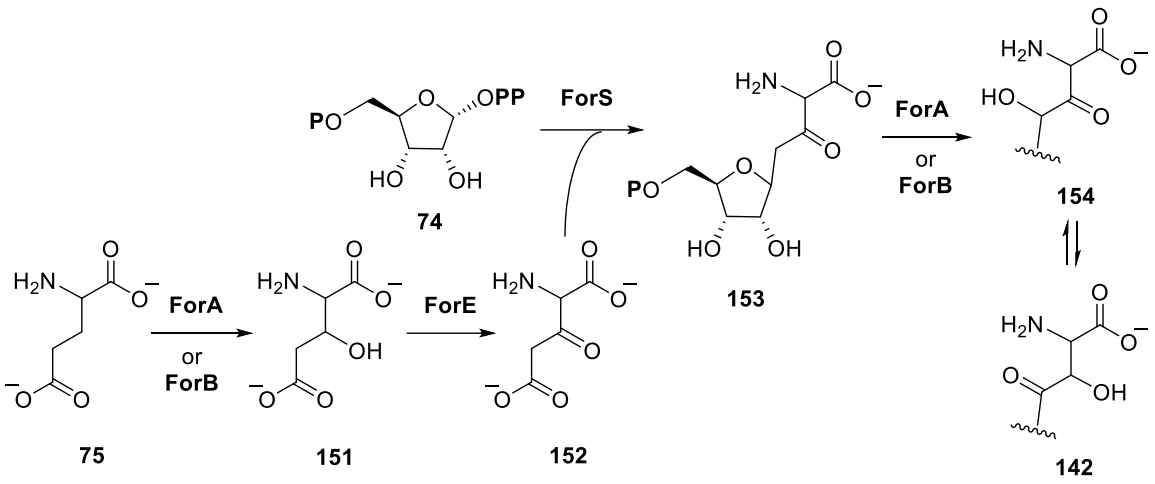


Figure 3-20. Alternative pathway for the formation of **142**.

Introduction of the ε-nitrogen of L-lysine (**76**) into the pyrazole ring is hypothesized to be catalyzed by a series of enzymes such as ForH, ForI, and ForJ, which appear to be a homolog of methionine–tRNA ligase, L-lysine 6-monooxygenase, and saccharopine dehydrogenase, respectively. *N*-Hydroxylysine (**143**) generated by ForI can be ligated with **142** by ForH to form **144**. Such an association of methionine–tRNA ligase with L-lysine 6-monooxygenase can be found in many bacteria producing peptide-derived antibiotics.[141] The following oxidation and hydrolysis catalyzed by ForJ may generate the intermediate **146** with the hydrazine moiety, which then be cyclized by ForK to afford **147**. The subsequent oxidation catalyzed by ForP or ForW, transamination by ForG, and iminium ion formation by ForP or ForW are proposed to form carboxyaminopyrazole ribonucleotide (**128**).

3.3.5 Biosynthesis of Coformycin

The formycin-producing bacteria also produce coformycin (**108**), a nucleoside antibiotic with a seven-membered heterocycle.[52,103] Coformycin (**108**) and its deoxyribose counterpart 2′-deoxycoformycin (**109**) are known to be derived from adenine nucleotide, and the expansion of the purine ring is from the insertion of one carbon from ribose (see Figure 3-2).[118,119] The early steps are proposed to be similar to histidine biosynthesis (**111**, **112** → **116**).[121] Thus, the genome of *S. antibioticus*, which produces **109**, was searched for homologous genes. The PCR-based screening was conducted using primers specifically designed based on the known conserved sequences of *hisG* and *hisA* in *Streptomyces* species, of which gene products catalyze the 1st and 4th step of histidine biosynthesis. The screening identified only one kind of *hisG* and *hisA* in *S. antibioticus*, supporting the idea that the biosynthesis of **108** and **109** utilizes the intermediate **116** from the histidine pathway.

The genes involved in the production of coformycin are expected to be located in close proximity to the formycin A gene cluster given their strongly associated physical properties. In the course of investigating the biosynthetic gene cluster for formycin A in *S. kaniharaensis*, it was noticed that one upstream gene adjacent to the gene cluster encodes an NAD(P)H-dependent reductase (*orf -2* in Figure 3-4). The early study using partially purified cell-free extract of *S. antibioticus* has shown that the reduction of 8-keto compounds (**119** and **120**) is a requisite step in the biosynthesis of coformycin, and the reduction reaction requires NADPH as the hydride donor.[120] The presence of an NAD(P)H-dependent reductase in the vicinity of formycin A gene cluster suggested that it could be responsible for the final reduction reaction to generate coformycin.

The reductase was expressed and purified as an *N*-terminal His_6-tagged protein, and the activity was assayed in the reverse reaction using the commercially available 2′-

deoxycoformycin (**109**) as a substrate. When 0.5 mM **109** was incubated with 0.4 μM enzyme in the presence of 1 mM $NADP^+$, the increase in the UV absorbance at 340 nm was observed indicating the production of NADPH concomitant with the oxidation of **109** (Figure 3-21). The reaction mixture containing NAD^+ instead of $NADP^+$ did not result in a significant increase in UV absorbance, consistent with the previous study.[120] The product was isolated by HPLC (Figure 3-22) and analyzed by ESI-MS (positive ion mode), which gave a clear signal at 289.1 *m*/*z*. This corresponded to the $[M + Na]^+$ ion of 8-ketodeoxycoformycin (**120**), suggesting that ORF -2 is involved in the reduction of 8-keto compounds (**119** and **120**). It is unknown whether the enzyme exhibits a comparable activity toward phosphorylated compounds (**118**). Further kinetic analysis is required to establish the newly identified ORF -2 as the pathway-specific reductase. While preliminary, these results support the identification of the reductase for the last step of coformycin biosynthesis, which corroborates the proposed pathway through the intermediary of 8-ketocoformycin (**119**).

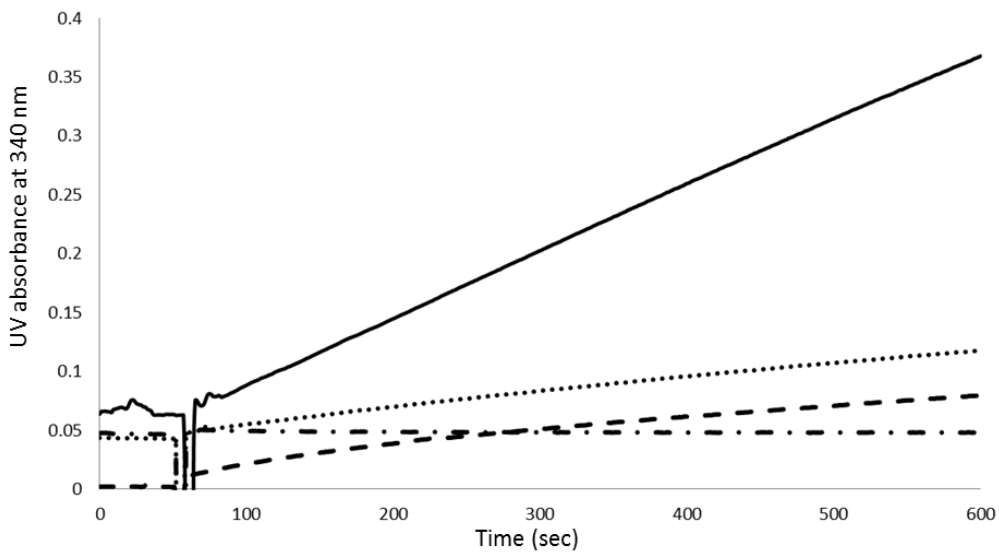


Figure 3-21. Activity assay of ORF -2. The enzyme was added at around 60 sec. The black line denotes the reaction with $NADP^+$, and the dotted line denotes the reaction with NAD^+ instead of $NADP^+$. The dashed line corresponds to the reaction containing neither $NADP^+$ nor NAD^+, and the dashed-dotted line corresponds to the reaction without substrate **109**.

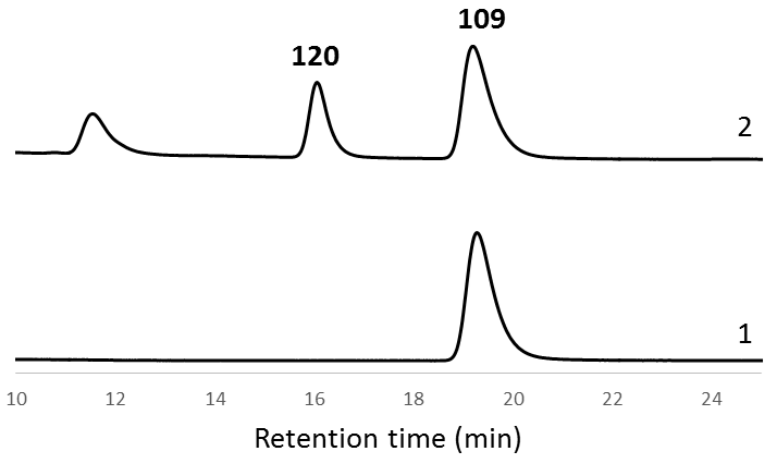


Figure 3-22. HPLC traces showing the reaction of ORF -2. Trace 1 is the reaction mixture containing 0.5 mM **109** and 1 mM $NADP^+$ before adding ORF -2, and trace 2 is the same reaction mixture as trace 1 incubated with 0.4 μM ORF -2 for 2 h.

3.4 CONCLUSION

Reported here is an investigation of the biosynthesis of formycin A (**51**) by sequencing the genome of *S. kaniharaensis*. The genomic analysis revealed the presence of a second set of *pur* homologous genes (*forV*, *forR*, *forC*, and *forD*) involved in the adenosine biosynthetic pathway. Given the structural resemblance of formycin A to adenosine, a similar pathway for the assembly of pyrazolopyrimidine moiety of formycin A was proposed, which includes carboxyaminopyrazole ribonucleotide (**128**) as a pathway intermediate. The putative biosynthetic gene cluster of formycin A was identified by screening a cosmid library to locate *pur*-like genes, and bioinformatics analysis of the entire gene cluster allowed a biosynthetic pathway of formycin A to be proposed. The enzymatic conversion of **128** to formycin A 5′-phosphate (**78**) except the formyl transfer step was demonstrated to be catalyzed by ForV, ForR, ForC, and ForD as proposed. Moreover, the specificity of ForD toward formycin B 5′-phosphate (**77**) over the closely related substrate IMP (**125**) was examined by a competition experiment, which showed that ForD exhibits 9-fold selectivity toward **77**. These observations strongly support the identification of a biosynthetic gene cluster specific for formycin A. Although no biochemical characterization has been realized yet, several enzymes that are proposed to catalyze early steps leading to the formation of intermediate **128** are being actively studied. In addition, a reductase encoded in close proximity to the formycin A gene cluster was found to be likely involved in the biosynthesis of coformycin (**108**), which is another nucleoside antibiotic possessing cooperative bioactivities. The discovery of a biosynthetic gene cluster for formycin A highlights a possible evolutionary link between a primary and a secondary metabolic pathway. A complete understanding of the biosynthetic pathway of formycin A and the enzymatic reaction mechanisms will shed light on the biosynthesis of

C-nucleosides, which also paves the way for repurposing of the pathway to develop new antibiotics.

References

1. Frey, P. A.; Hegeman, A. D.; Ruzicka, F. J. The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63–88.

2. Frey, P. A. Travels with carbon-centered radicals. 5′-Deoxyadenosine and 5′-deoxyadenosine-5′-yl in radical enzymology. *Acc. Chem. Res.* **2014**, *47*, 540–549.

3. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **2014**, *114*, 4229−4317.

4. Thorson, J. S.; Hosted Jr., T. J.; Jiang, J.; Biggins, J. B.; Ahlert, J. Nature's Carbohydrate Chemists The Enzymatic Glycosylation of Bioactive Bacterial Metabolites. *Curr. Org. Chem.* **2001**, *5*, 139−167.

5. Ruszczycky, M. W.; Ogasawara, Y.; Liu, H.-w. Radical SAM enzymes in the biosynthesis of sugar-containing natural products. *Biochim. Biophys. Acta.* **2012**, *1824*, 1231−1244.

6. Llewellyn, N. M.; Spencer, J. B. Biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *Nat. Prod. Rep.* **2006**, *23*, 864–874.

7. Ota, Y.; Tamegai, H.; Kudo, F.; Kuriki, H.; Koike-Takeshita, A.; Eguchi, T. Butirosin biosynthetic gene cluster from *Bacillus circulans*. *J. Antibiot.* **2000**, *53*, 1158–1167.

8. Tamegai, H.; Nango, E.; Kuwahara, M.; Yamamoto, H.; Ota, Y.; Kuriki, H.; Eguchi, T.; Kakinuma, K. Identification of L-glutamine:2-deoxy-scyllo-inososе amino-transferase required for the biosynthesis of butirosin in *Bacillus circulans*. *J. Antibiot.* **2002**, *55*, 707–714.

9. Yokoyama, K.; Numakura, M.; Kudo, F.; Ohmori, D.; Eguchi, T. Characterization and mechanistic study of a radical SAM dehydrogenase in the biosynthesis of butirosin. *J. Am. Chem. Soc.* **2007**, *129*, 15147–15155.

10. Yokoyama, K.; Ohmori, D.; Kudo, F.; Eguchi, T. Mechanistic study on the reaction of a radical SAM dehydrogenase BtrN by electronic paramagnetic resonance spectroscopy, *Biochemistry* **2008**, *47*, 8950–8960.

11. Grove, T. L.; Ahlum, J. H.; Sharma, P.; Krebs, C.; Booker, S. J. A consensus mechanism for radical SAM-dependent dehydrogenation? BtrN contains two [4Fe–4S] clusters. *Biochemistry* **2010**, *49*, 3783–3785.

12. Goldman, P. J.; Grove, T. L.; Booker, S. J.; Drennan, C. L. X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12111–12116.

13. Maiocco, S. J.; Grove, T. L.; Booker, S. J.; Elliott, S. J. Electrochemical Resolution of the [4Fe-4S] Centers of the AdoMet Radical Enzyme BtrN: Evidence of Proton Coupling and an Unusual, Low-Potential Auxiliary Cluster. *J. Am. Chem. Soc.* **2015**, *137*, 8664−8667.

14. Booker, S. J. Anaerobic functionalization of unactivated C−H bonds. *Curr. Opin. Chem. Biol.* **2009**, *13*, 58–73.

15. Unwin, J.; Standage, S.; Alexander, D.; Hosted Jr., T.; Horan, A. C.; Wellington, E.M.H. Gene cluster in *Micromonospora echinospora* ATCC15835 for the biosynthesis of the gentamicin C complex. *J. Antibiot.* **2004**, *57*, 436–445.

16. Huang, C.; Huang, F.; Moison, E.; Guo, J.; Jian, X.; Duan, X.; Deng, Z.; Leadlay, P. F.; Sun, Y. Delineating the Biosynthesis of Gentamicin X_2, the Common Precursor of the Gentamicin C Antibiotic Complex. *Chem. Biol.* **2015**, *22*, 251–261.

17. Kim, H. J.; McCarty, R. M.; Ogasawara, Y.; Liu, Y.-n.; Mansoorabadi, S. O.; LeVieux, J.; Liu, H.-w. GenK-Catalyzed C-6′ Methylation in the Biosynthesis of Gentamicin: Isolation and Characterization of a Cobalamin-Dependent Radical SAM Enzyme. *J. Am. Chem. Soc.* **2013**, *135*, 8093−8096.

18. Kuzuyama, T.; Seki, T.; Dairi, T.; Hidaka, T.; Seto, H. Nucleotide sequence of fortimicin KL1 methyltransferase gene isolated from *Micromonospora olivasterospora*, and comparison of its deduced amino acid sequence with those of methyltransferases involved in the biosynthesis of bialaphos and fosfomycin. *J. Antibiot.* **1995**, *48*, 1191−1193.

19. Ostash, B.; Doud, E. H.; Lin, C.; Ostash, I.; Perlstein, D. L.; Fuse, S.; Wolpert, M.; Kahne, D.; Walker, S. Complete Characterization of the Seventeen Step Moenomycin Biosynthetic Pathway. *Biochemistry* **2009**, *48*, 8830–8841.

20. Nagahata, T.; Ueda, K.; Tsurimoto, T.; Chisaka, O.; Matsubara, K. Anti-hepatitis B virus activities of purine derivatives of oxetanocin A. *J. Antibiot.* **1989**, *42*, 644–646.

21. Morita, M.; Tomita, K.; Ishizawa, M.; Takagi, K.; Kawamura, F.; Takahashi, H.; Morino, T. Cloning of oxetanocin A biosynthetic and resistance genes that reside on a plasmid of *Bacillus megaterium* strain NK84-0128, *Biosci. Biotechnol. Biochem.* **1999**, *63*, 563–566.

22. Morita, M.; Tomita, K.; Ishizawa, M.; Takagi, K.; Kawamura, F.; Takahashi, H.; Morino, T. Molecular cloning of oxetanocin A biosynthetic and resistance genes which reside on a plasmid of *Bacillus megaterium* strain NK84-0128, Unpublished direct submission to NCBI Data Base, http://www.ncbi.nlm.nih.gov, **2007**, Accession No. AB005787.

23. Kimura, K.-I.; Bugg, T. D. H. Recent advances in antimicrobial nucleoside antibiotics targeting cell wall biosynthesis. *Nat. Prod. Rep.* **2003**, *20*, 252–273.

24. Tsvetanova, B. C.; Kiemle, D. J.; Price, N. P. J. Biosynthesis of tunicamycin and metabolic origin of the 11-carbon dialdose sugar, tunicamine. *J. Biol. Chem.* **2002**, *277*, 35289–35296.

25. Wyszynski, F. J.; Hesketh, A. R.; Bibb, M. J.; Davis, B. G. Dissecting tunicamycin biosynthesis by genome mining: cloning and heterologous expression of a minimal gene cluster. *Chem. Sci.* **2010**, *1*, 581–589.

26. Chen, W.; Qu, D.; Zhai, L.; Tao, M.; Wang, Y.; Lin, S.; Price, N. P. J.; Deng, Z. Characterization of the tunicamycin gene cluster unveiling unique steps involved in its biosynthesis. *Protein Cell* **2010**, *1*, 1093–1105.

27. Wyszynski1, F. J.; Lee, S. S.; Yabe, T.; Wang, H.; Gomez-Escribano, J. P.; Bibb, M. J.; Lee, S. J.; Davies, G. J.; Davis, B. G. Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nat. Chem.* **2012**, *4*, 539–546.

28. Yamaguchi, H.; Yamamoto, C.; Tanaka, N. Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells. *J. Biochem.* **1965**, *57*, 667–677.

29. Iwasa, T.; Suetomi, K.; Kusaka, T. Taxonomic study and fermentation of producing organism and antimicrobial activity of mildiomycin. *J. Antibiot.* **1978**, *31*, 511–518.

30. Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the Mildiomycin Biosynthesis Gene Cluster in *Streptoverticillum remofaciens* ZJU5119 and Characterization of MilC, a Hydroxymethyl cytosylglucuronic Acid Synthase. *ChemBioChem* **2012**, *13*, 1613–1621.

31. Feng, J.; Wu, J.; Dai, N.; Lin, S.; Xu, H. H.; Deng, Z.; He, X. Discovery and Characterization of BlsE, a Radical *S*-Adenosyl-L-methionine Decarboxylase Involved in the Blasticidin S Biosynthetic Pathway. *PLoS ONE* **2013**, *8*, e68545.

32. Wright, D. E. The orthosomycins, a new family of antibiotics. *Tetrahedron* **1979**, *35*, 1207–1237.

33. Weitnauer, G.; Mühlenweg, A.; Trefzer, A.; Hoffmeister, D.; Süβmuth, R. D.; Jung, G.; Welzel, K.; Vente, A.; Girreser, U.; Bechthold, A. Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the *avi* biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chem. Biol.* **2001**, *8*, 569–581.

34. Boll, R.; Hofmann, C.; Heitmann, B.; Hauser, G.; Glaser, S.; Koslowski, T.; Friedrich, T.; Bechthold, A. The Active Conformation of Avilamycin A Is Conferred by AviX12, a Radical AdoMet Enzyme. *J. Biol. Chem.* **2006**, *281*, 14756–14763.

35. Schlünzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001**, 413, 814–821.

36. Xue, Y.; Zhao, L.; Liu, H.-w.; Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12111–12116.

37. Zhao, L.; Borisova, S.; Yeung, S-M.; Liu, H.-w. Study of C-4 Deoxygenation in the Biosynthesis of Desosamine: Evidence Implicating a Novel Mechanism. *J. Am. Chem. Soc.* **2001**, *123*, 7909–7910.

38. Szu, P.-h.; He, X.; Zhao, L.; Liu, H.-w. Biosynthesis of TDP-D-Desosamine: Identification of a Strategy for C4 Deoxygenation. *Angew. Chem. Int. Ed.* **2005**, *44*, 6742–6746.

39. Borisova, S. A.; Zhao, L.; Melancon, III, C. E.; Kao, C.-L.; Liu, H.-w. Characterization of the Glycosyltransferase Activity of DesVII: Analysis of and Implications for the Biosynthesis of Macrolide Antibiotics. *J. Am. Chem. Soc.* **2004**, *126*, 6534–6535.

40. Szu, P.-H.; Ruszczycky, M. W.; Choi, S.-h.; Liu, H.-w. Characterization and mechanistic studies of DesII: a radical *S*-adenosyl-L-methionine enzyme involved in the biosynthesis of TDP-D-desosamine. *J. Am. Chem. Soc.* **2009**, *131*, 14030–14042.

41. Ruszczycky, M. W.; Choi, S.-h.; Liu, H.-w. Stoichiometry of the Redox Neutral Deamination and Oxidative Dehydrogenation Reactions Catalyzed by the Radical SAM Enzyme DesII. *J. Am. Chem. Soc.* **2010**, *132*, 2359–2369.

42. Ruszczycky, M. W.; Choi, S.-h.; Mansoorabadi, S. O.; Liu, H.-w. Mechanistic Studies of the Radical *S*-Adenosyl-L-methionine Enzyme DesII: EPR Characterization of a Radical Intermediate Generated During Its Catalyzed Dehydrogenation of TDP-D-Quinovose. *J. Am. Chem. Soc.* **2011**, *133*, 7292–7295.

43. Ruszczycky, M. W.; Choi, S.-h.; Liu, H.-w. EPR-kinetic isotope effect study of the mechanism of radical-mediated dehydrogenation of an alcohol by the radical SAM enzyme DesII. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 2088–2093.

44. Hayon, E.; Simic, M. Acid–base properties of free radicals in solution. *Acc. Chem. Res.* **1974**, *7*, 114–121.

45. Rao, P. S.; Hayon, E. Correlation between ionization constants of organic free radicals and electrochemical properties of parent compounds. *Anal. Chem.* **1976**, *48*, 564–568.

46. Bandarian, V.; Reed, G. H. Ethanolamine Ammonia-Lyase. In *Chemistry and Biochemistry of B_{12}*; Banerjee, R., Ed.; Wiley-Interscience: New York, 1999; pp 811–833.

47. Semialjac, M.; Schwarz, H. Computational Exploration of Rearrangements Related to the Vitamin B_{12}-Dependent Ethanolamine Ammonia Lyase Catalyzed Transformation. *J. Am. Chem. Soc.* **2002**, *124*, 8974–8983.

48. Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J. Antibiot.* **1988**, *41*, 1711–1739.

49. Winn, M,; Goss, R. J. M.; Kimura, K.-i.; Bugg, T. D. H. Antimicrobial nucleoside antibiotics targeting cell wall assembly: Recent advances in structure–function studies and nucleoside biosynthesis. *Nat. Prod. Rep.* **2010**, *27*, 279–304.

50. De Clercq, E. C-Nucleosides To Be Revisited. *J. Med. Chem.* **2016**, *59*, 2301−2311.

51. Hori, M.; Ito, E.; Takita, T.; Koyama, G.; Takeuchi, T.; Umezawa H. A new antibiotic, formycin. *J. Antibiotics, Ser. A* **1964**, *17*, 96–99.

52. Suhadolnik, R. J. *Nucleoside Antibiotics*; Wiley-Interscience, New York, N. Y., 1970.

53. Koyama, G.; Umezawa H. Formycin B and its relation to formycin. *J. Antibiotics, Ser. A* **1965**, *18*, 175–177.

54. Williams, R. H.; Hoehn, M. M. U.S. Patent 3,674,774, 1972.

55. Nishimura, H.; Mayama, M.; Komatsu, Y.; Kato, H.; Shimaoka, N.; Tanaka, Y. Showdomycin, a new antibiotic from a *Streptomyces* sp. *J. Antibiotics, Ser. A* **1964**, *17*, 148–155.

56. Kusakabe, Y.; Nagatsu, J.; Shibuya, M.; Kawaguchi, O.; Hirose, C. Minimycin, a new antibiotic. *J. Antibiotics, Ser. A* **1972**, *25*, 44–47.

57. Mueller, E. G.; Ferré-D'Amaré, A. R. In *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*; Grosjean, H., Ed.; Landes Bioscience: Austin, TX, 2009; p 3633.

58. Veerareddygari, G. R.; Singh, S. K.; Mueller, E. G. The Pseudouridine Synthases Proceed through a Glycal Intermediate. *J. Am. Chem. Soc.* **2016**, *138*, 7852−7855.

59. Huang, L.; Pookanjanatavip, M,; Gu, X.; Santi, D. V. A Conserved Aspartate of tRNA Pseudouridine Synthase Is Essential for Activity and a Probable Nucleophilic Catalyst. *Biochemistry* **1998**, *37*, 344−351.

60. Sakata, K.; Sakurai, A.; Tamura, S. Isolation of Novel Antifungal Antibiotics, Ezomycins A_1, A_2, B_1 and B_2. *Agr. Biol. Chem.* **1974**, *38*, 1883−1890.

61. Sakata, K.; Sakurai, A.; Tamura, S. Structures of Ezomycins A_l and A_2. *Agr. Biol. Chem.* **1976**, *40*, 1993−1999.

62. Sakata, K.; Sakurai, A.; Tamura, S. Structures of Ezomycins B_1, B_2, C_1, C_2, D_1 and D_2. *Agr. Biol. Chem.* **1977**, *41*, 2033−2039.

63. Hanessian, S.; Marcotte, S.; Machaalani, R.; Huang, G. Total Synthesis and Structural Confirmation of Malayamycin A: A Novel Bicyclic *C*-Nucleoside from *Streptomyces malaysiensis*. *Org. Lett.* **2003**, *5*, 4277−4280.

64. Sato, T.; Hirasawa, K.; Uzawa, J.; Inaba, T.; Isono, K. Biosynthesis of octosyl acid A: Incorporation of C-13 labeled glucose. *Tetrahedron Lett.* **1979**, *20*, 3441–3444.

65. Lilla, E. A.; Yokoyama, K. Carbon extension in peptidylnucleoside biosynthesis by radical SAM enzymes. *Nat. Chem. Biol.* **2016**, *12*, 905–907.

66. Sawa, T.; Fukagawa, Y.; Homma, I.; Wakashiro, T.; Takeuchi, T.; Hori, M.; Komai, T. Metabolic conversion of formycin B to formycin A and to oxoformycin B in *Nocardia interforma*. *J. Antibiot.* **1968**, *21*, 334–339.

67. Kunimoto, T.; Sawa, T.; Wakashiro, T.; Hori, M.; Umezawa, H. Biosynthesis of the formycin family. *J. Antibiot.* **1971**, *24*, 253–258.

68. Ochi, K.; Iwamoto, S.; Hayase, E.; Yashima, S.; Okami, Y. Biosynthesis of formycin. Role of certain amino acids in formycin biosynthesis. *J. Antibiot.* **1974**, *27*, 909–916.

69. Ochi, K.; Yashima, S.; Eguchi, Y. Matsushita, K. Biosynthesis of formycin. Incorporation and distribution of ^{13}C-, ^{14}C-, and ^{15}N-labeled compounds into formycin. *J. Biol. Chem.* **1979**, *254*, 8819–8824.

70. Buchanan, J. G.; Hamblin, M. R.; Sood, G. R.; Wightman, R. H. The Biosynthesis of Pyrazofurin and Formycin. *J.C.S. Chem. Comm.* **1980**, 917–918.

71. Buchanan, J. M.; Hartman, S. C. Enzymic reactions in the synthesis of the purines. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1959**, *21*, 199–261.

72. Elstner, E. F.; Suhadolnik, R. J. Nucleoside Antibiotics. Asymmetric Incorporation of Glutamic Acid and Acetate into the Maleimide Ring of Showdomycin by *Streptomyces showdoensis*. *Biochemistry* **1972**, *11*, 2578–2584.

73. Isono, K.; Suhadolnik, R. J. Biosynthesis of the C-nucleoside, minimycin: asymmetric incorporation of glutamate and acetate into the oxazine ring. *J. Antibiot.* **1977**, *30*, 272–273.

74. Hultin, P. G. Bioactive C-glycosides from bacterial secondary metabolism. *Curr. Topics Med. Chem.* **2005**, *5*, 1299–1331.

75. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555.

76. Durr, C.; Hoffmeister, D.; Wohlert, S. E.; Ichinose, K.; Weber, M.; Von Mulert, U.; Thorson, J. S.; Bechthold, A. The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers. *Angew. Chem. Int. Ed.* **2004**, *43*, 2962–2965.

77. Baig, I.; Kharel, M.; Kobylyanskyy, A.; Zhu, L.; Rebets, Y.; Ostash, B.; Luzhetskyy, A.; Bechthold, A.; Fedorenko, V. A.; Rohr, J. On the acceptor substrate of C-glycosyltransferase UrdGT2: three prejadomycin C-Glycosides from an engineered mutant of *Streptomyces globisporus* 1912 Δ*lndE*(*urdGT2*). *Angew. Chem. Int. Ed.* **2006**, *45*, 7842–7846.

78. Brazier-Hicks, M.; Evans, K. M.; Gershater, M. C.; Puschmann, H.; Steel, P. G.; Edwards, R. The *C*-Glycosylation of Flavonoids in Cereals. *J. Biol. Chem.* **2009**, *284*, 17926–17934.

79. Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. *In vitro* characterization of IroB, a pathogen-associated *C*-glycosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 571–576.

80. Thibodeaux, C. J.; Melançon, C. E.; Liu, H.-w. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **2007**, *446*, 1008–1016.

81. Chen, D.; Chen, R.; Wang, R.; Li, J.; Xie, K.; Bian, C.; Sun, L.; Zhang, X.; Liu, J.; Yang, L.; Ye, F.; Yu, X.; Dai, J. Probing the Catalytic Promiscuity of a Regio- and Stereospecific C-Glycosyltransferase from *Mangifera indica*. *Angew. Chem. Int. Ed.* **2015**, *54*, 12678–12682.

82. Härle, J.; Gunther, S.; Lauinger, B.; Weber, M.; Kammerer, B.; Zechel, D. L.; Luzhetskyy, A.; Bechthold, A. Rational Design of an Aryl-C-Glycoside Catalyst from a Natural Product O-Glycosyltransferase. *Chem. Biol.* **2011**, *18*, 520–530.

83. Gutmann, A.; Nidetzky, B. Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs. *Angew. Chem. Int. Ed.* **2012**, *51*, 12879–12883.

84. Toraya, T. Radical Catalysis in Coenzyme B_{12}-Dependent Isomerization (Eliminating) Reactions. *Chem. Rev.* **2003**, *103*, 2095−2127.

85. Bansal, K. M.; Gratzel, M.; Henglein, A.; Janata, E. Polarographic and optical absorption studies of radicals produced in the pulse radiolysis of aqueous solutions of ethylene glycol. *J. Phys. Chem.* **1973**, *77*, 16–19.

86. Steenken, S. Oxidation of phenolates and phenylenediamines by 2-alkononyl radicals produced from 1,2-dihydroxy- and 1-hydroxy-2-alkoxyalkyl radicals. *J. Phys. Chem.* **1979**, *83*, 595–599.

87. Steenken, S.; Davies, M. J.; Gilbert, B. C. Pulse radiolysis and election spin resonance studies of the dehydration of radicals from 1,2-diols and related compounds. *J. Chem. Soc. Perkin Trans. 2* **1986**, 1003–1010.

88. Ruszczycky, M. W.; Liu, H.-w. Mechanistic Enzymology of the Radical SAM Enzyme DesII. *Isr. J. Chem.* **2015**, *55*, 315–324.

89. Ko, Y.; Ruszczycky, M. W.; Choi, S.-H.; Liu, H.-w. Mechanistic studies of the radical *S*-adenosylmethionine enzyme DesII with TDP-D-fucose. *Angew. Chem. Int. Ed.* **2015**, *54*, 860–863.

90. Pfoestl, A., Hofinger, A., Kosma, P., Messner, P. Biosynthesis of dTDP-3-acetamido-3,6-dideoxy-*a*-D-galactose in *Aneurinibacillus thermoaerophilus* L4290-91T. *J. Biol. Chem.* **2003**, *278*, 26410–26417.

91. Romana, L. K.; Santiago, F. S.; Reeves, P. R. High level expression and purification of dthymidine diphospho-D-glucose 4,6-dehydratase (rfbB) from *Salmonella* serovar typhimurium lt2. *Biochm. Biophys. Res. Comm.* **1991**, *174*, 846–852.

92. Melancon, III, C. E.; Hong, L.; White, J. A.; Liu, Y.-n.; Liu, H.-w. Characterization of TDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase from the D-mycaminose biosynthetic pathway of *Streptomyces fradiae*: In vitro activity and substrate specificity studies. *Biochemistry* **2007**, *46*, 577–590.

93. Park, J.; Tai, J.; Roessner, C. A.; Scott, A. I. Enzymatic synthesis of *S*-adenosyl-L-methionine on the preparative scale. *Bioorg. Med. Chem.* **1996**, *4*, 2179–2185.

94. Iwig, D. F.; Booker, S. J. Insight into the polar reactivity of the onium chalcogen analogues of *S*-adenosyl-L-methionine. *Biochemistry* **2004**, *43*, 13496–13509.

95. Choi, S. H. *Synthetic approaches to investigate the chemical mechanism in the biosynthesis of natural products*. Ph.D. Thesis, University of Texas at Austin, Austin, Texas.

96. Elling, L.; Rupprath, C.; Günther, N.; Römer, U.; Verseck, S.; Weingarten, P.; Dräger, G.; Kirschning, A.; Piepersberg, W. An enzyme module system for the synthesis of dTDP-activated deoxysugars from dTMP and sucrose. *ChemBioChem* **2005**, *6*, 1423–1430.

97. Chandor, A.; Berteau, O.; Douki, T.; Gasparutto, D.; Sanakis, Y.; Ollagnierde-Choudens, S.; Atta, M.; Fontecave, M. Dinucleotide Spore Photoproduct, a Minimal Substrate of the DNA Repair Spore Photoproduct Lyase Enzyme from *Bacillus subtilis*. *J. Biol. Chem.* **2006**, *281*, 26922–26931.

98. Chandor-Proust, A.; Berteau, O.; Douki, T.; Gasparutto, D.; Ollagnier-de-Choudens, S.; Fontecave, M.; Atta, M. DNA Repair and Free Radicals, New Insights into the Mechanism of Spore Photoproduct Lyase Revealed by Single Amino Acid Substitution. *J. Biol. Chem.* **2008**, *283*, 36361–36368.

99. Zhang, Y.; Zhu, X.; Torelli, A.T.; Lee, M.; Dzikovski, B.; Koralewski, R.M.; Wang, E.; Freed, J.; Krebs, C.; Ealick, S.E.; Lin, H. Diphthamide biosynthesis requires an organic radical generated by an iron–sulphur enzyme. *Nature* **2010**, *465*, 891–898.

100. Melancon, III, C. E.; Liu, H.-w. Engineered Biosynthesis of Macrolide Derivatives Bearing the Non-Natural Deoxysugars 4-*epi*-D-Mycaminose and 3-*N*-Monomethylamino-3-Deoxy-D-Fucose. *J. Am. Chem. Soc.* **2007**, *129*, 4896–4897.

101. Duschene, K. S.; Veneziano, S. E.; Silver, S. C.; Broderick, J. B. Control of radical chemistry in the AdoMet radical enzymes. *Curr. Opin. Chem. Biol.* **2009**, *13*, 74–83.

102. Lin, G.-M.; Choi, S.-H.; Ruszczycky, M. W.; Liu, H.-w. Mechanistic Investigation of the Radical *S*-Adenosyl-L-methionine Enzyme DesII Using Fluorinated Analogues. *J. Am. Chem. Soc.* **2015**, *137*, 4964–4967.

103. Nakamura, H.; Koyama, G.; Iitaka, Y.; Ono, M.; Yagiawa, N. Structure of coformycin, an unusual nucleoside of microbial origin. *J. Am. Chem. Soc.* **1974**, *96*, 4327–4328.

104. Sawa, T.; Fukagawa, Y.; Homma, I.; Takeuchi, T.; Umezawa, H. Mode of inhibition of coformycin on adenosine deaminase. *J. Antibiotics, Ser. A* **1967**, *20*, 227–231.

105. Woo, P. W. K.; Dion, H. W.; Lange, S. M.; Dahl, L. F.; Durham, L. J. A novel adenosine and ara-a deaminase inhibitor, (R)-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-3,6,7,8-tetrahydroimidazo[4,5-*d*][1,3]diazepin-8-ol. *J. Heterocycl. Chem.* **1974**, *11*, 641–643.

106. Jain, P.; Pemmaraju, N.; Ravandi, F. Update on the biology and treatment options for hairy cell leukemia. *Curr. Treat. Options Oncol.* **2014**, *15*, 187–209.

107. Suzuki, M.; Okuda, T.; Shiraki, K. Synergistic antiviral activity of acyclovir and vidarabine against herpes simplex virus types 1 and 2 and varicella-zoster virus. *Antiviral Res.* **2006**, *72*, 157–161.

108. Schwartz, P. M.; Shipman, C. Jr.; Drach, J. C. Antiviral activity of arabinosyladenine and arabinosylhypoxanthine in herpes simplex virus-infected KB cells: selective inhibition of viral deoxyribonucleic acid synthesis in the presence of an adenosine deaminase inhibitor. *Antimicrob. Agents Chemother.* **1976**, *10*, 64–74.

109. Hershfield, M. S.; Kredich, N. M.; Koller, C. A.; Mitchell, B. S.; Kurtzberg, J.; Kinney, T. R.; Falletta, J. M. S-adenosylhomocysteine catabolism and basis for acquired resistance during treatment of T-cell acute lymphoblastic leukemia with 2'-deoxycoformycin alone and in combination with 9-beta-D-arabinofuranosyladenine. *Cancer Res.* **1983**, *43*, 3451–3458.

110. Bzowska, A.; Kulikowska, E.; Shugar, D. Formycins A and B and some analogues: selective inhibitors of bacterial (*Escherichia coli*) purine nucleoside phosphorylase. *Biochim. Biophys. Act.* **1992**, *1120*, 239–247.

111. Kierdaszuk, B.; Modrak-Wójcik, A.; Wierzchowski, J.; Shugar, D. Formycin A and its *N*-methyl analogues, specific inhibitors of *E. coli* purine nucleoside phosphorylase (PNP): induced tautomeric shifts on binding to enzyme, and enzyme→ligand fluorescence resonance energy transfer. *Biochim. Biophys. Act.* **2000**, *1476*, 109–128.

112. Long, M. C.; Parker, W. B. Structure–activity relationship for nucleoside analogs as inhibitors or substrates of adenosine kinase from *Mycobacterium tuberculosis*. I. Modifications to the adenine moiety. *Biochem. Pharmacol.* **2006**, *71*, 1671–1682.

113. Cornell, K. A.; Swarts, W. E.; Barry, R. D.; Riscoe, M. K. Characterization of recombinant *Eschericha coli* 5′-methylthioadenosine/*S*-adenosylhomocysteine nucleosidase: analysis of enzymatic activity and substrate specificity. *Biochem. Biophys. Res. Commun.* **1996**, *228*, 724–732.

114. Ishizuka, M.; Takeuchi, T.; Nitta, K.; Koyama, G.; Hori, M.; Umezawa, H. Antitumor activities of formycin and labilomycin. *J. Antibiotics. Ser. A.* **1964**, *17*, 124–126.

115. Ishizuka, M.; Sawa, T.; Hori, S.; Takayama, H.; Takeuchi, T.; Umezawa, H. Biological studies on formycin and formycin B. *J. Antibiot.* **1968**, *21*, 5–12.

116. Takeuchi, T.; Iwanaga, J.; Aoyagi, T.; Umezawa, H. Antiviral effect of formycin and formycin B. *J. Antibiotics. Ser. A.* **1966**, *19*, 286–287.

117. Dapp, M. J.; Bonnac, L.; Patterson, S. E.; Mansky, L. M. Discovery of Novel Ribonucleoside Analogs with Activity against Human Immunodeficiency Virus Type 1. *J. Virol.* **2014**, *88*, 354–363.

118. Hanvey, J. C.; Hardman, J. K.; Suhadolnik, R. J.; Baker, D. C. Evidence for the Conversion of Adenosine to 2′-Deoxycoformycin by *Streptomyces antibioticus*. *Biochemistry* **1984**, *23*, 904–907.

119. Hanvey, J. C.; Hawkins, E. S.; Tunac, J. B.; Dechter, J. J.; Baker, D. C.; Suhadolnik, R. J. Biosynthesis of 2′-Deoxycoformycin: Evidence for Ring Expansion of the Adenine Moiety of Adenosine to a Tetrahydroimidazo[4,5-*d*][1,3]diazepine System. *Biochemistry* **1987**, *26*, 5636–5641.

120. Hanvey, J. C.; Hawkins, E. S.; Baker, D. C.; Suhadolnik, R. J. 8-Ketodeoxycoformycin and 8-Ketocoformycin as Intermediates in the Biosynthesis of 2′-Deoxycoformycin and Coformycin. *Biochemistry* **1988**, *27*, 5790–5795.

121. Smith, D. W. E.; Ames, B. N. Intermediates in the Early Steps of Histidine Biosynthesis. *J. Biol. Chem.* **1964**, *239*, 1848–1855.

122. Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual;* Cold Spring Harbor Laboratory Press: New York, 2001.

123. Chevreux, B.; Wetter, T.; Suhai, S. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB). **1999**, *99*, 45–56.

124. Delcher, A. L.; Bratke, K. A.; Powers, E. C.; Salzberg, S. L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **2007**, *23*, 673–679.

125. Lieberman, I. Enzymatic synthesis of adenosine-5'-phosphate from inosine-5'-phosphate. *J. Biol. Chem.* **1956**, *223*, 327–339.

126. Lipps, G.; Krauss, G. Adenylosuccinate synthase from *Saccharomyces cerevisiae*: homologous overexpression, purification and characterization of the recombinant protein. *Biochem. J.* **1999**, *341*, 537–543.

127. Rabinowitz, J. C. Preparation and properties of 5,10-methenyltetrahydrofolic acid and 10-formyltetrahydrofolic acid. **1963**, *6*, *Methods Enzymol.* 814–815.

128. Meyer, E.; Leonard, N. J.; Bhat, B.; Stubbe, J.; Smith, J. M. Purification and characterization of the purE, purK, and purC gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway. *Biochemistry* **1992**, *31*, 5022–5032.

129. Rayl, E. A.; Moroson, B. A.; Beardsley, G. P. The human purH gene product, 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase / IMP cyclohydrolase. Cloning, sequencing, expression, purification, kinetic analysis, and domain mapping. *J. Biol. Chem.* **1996**, *271*, 2225–2233.

130. Bigeleisen, J.; Wolfsberg, M. Theoretical and Experimental Aspects of Isotope Effects in Chemical Kinetics. *Adv. Phys. Chem.* **1958**, *1*, 15–76.

131. Grindey, G. B.; Nichol, C. A. Micro procedure for determination of pyrophosphate and orthophosphate. *Anal. Biochem.* **1970**, *33*, 114–119.

132. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. Practical Streptomyces Genetics; The John Innes Foundation: Norwich, UK, 2000; pp 161–171.

133. Zalkin, H.; Dixon, J. E. De novo purine nucleotide biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.* **1992**, *42*, 259–287.

134. Zhang, Y.; Morar, M.; Ealick, S. E. Structural biology of the purine biosynthetic pathway. *Cell. Mol. Life Sci.* **2008**, *65*, 3699–3724.

135. Medema, M. H.; Fischbach, M. A. Computational approaches to natural product discovery. *Nat. Chem. Biol.* **2015**, *11*, 639–648.

136. Ju, K.-S.; Gao, J.; Doroghazi, J. R.; Wang, K.-K. A.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cio-ni, J. P.; Evans, B. S.; Hirota, R.; Labeda, D. P.; van der Donk, W. A.; Metcalf, W. W. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12175–12180.

137. Penn, K.; Jenkins, C.; Nett, M.; Udwary, D. W.; Gontang, E. A.; McGlinchey, R. P.; Foster, B.; Lapidus, A.; Podell, S.; Allen, E. E.; Moore, B. S. Jensen. P. R. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J.* **2009**, *3*, 1193–1203.

138. Madduri, K.; Stuttard, C.; Vining, L. C. Lysine Catabolism in *Streptomyces* spp. Is Primarily through Cadaverine: β-Lactam Producers Also Make α-Aminoadipate. *J. Bacteriol.* **1989**, 299–302.

139. Dumitru, R. V.; Ragsdale, S. W. Mechanism of 4-(*β*-D-Ribofuranosyl)aminobenzene 5′-Phosphate Synthase, a Key Enzyme in the Methanopterin Biosynthetic Pathway. *J. Biol. Chem.* **2004**, *279*, 39389–39395.

140. White, R. H. The Conversion of a Phenol to an Aniline Occurs in the Biochemical Formation of the 1-(4-Aminophenyl)-1-deoxy-D-ribitol Moiety in Methanopterin. *Biochemistry* **2011***, 50*, 6041–6052.

141. Aravind, L.; de Souza, R. F.; Iyer, L. M. Predicted class-I aminoacyl tRNA synthetase-like proteins in non-ribosomal peptide synthesis. *Biol. Direct* **2010**, *5*, 48.